A Framework for Evaluating Proposed Technologies for Next-Generation Wireless Systems

By

WILLIAM JONES

Industrial Doctorate Centre in Systems
UNIVERSITY OF BRISTOL

A dissertation submitted to the University of Bristol in accordance with the requirements of the degree of ENGINEERING DOCTORATE in the Faculty of Engineering.

AUGUST, 2018

Word count: 64,715

Abstract

This thesis presents a framework for evaluating proposed technologies for next-generation wireless systems, using systems modelling approaches. First, the socio-economic system is explored addressing the challenging question of how to develop a strategy for research investment in the complex development space of Fifth Generation (5G) era technologies. By the application of Problem Structuring Methods, and focusing on developing a clearer understanding of the industry landscape, a methodology for strategic decision making is proposed. The approach is used to identify key areas of wireless technology research for the 5G era.

Subsequently, identified key areas of wireless technology including, full-duplex, beamforming, clear channel assessment and transmission power adaptation are explored in single and multi-hop wireless networks. A novel conceptual simulation modelling methodology is proposed and applied to investigate the performance impact of these technologies when implemented in the context of Carrier Sense Multiple Access with Collision Avoidance wireless networks. The methodology is designed to aid researchers in the environment of a corporate research and development lab with the goal of developing innovations and intellectual property that can bring commercial success. Whilst each technology is capable in principle of improving system performance, often the gain is limited when implementing in a network environment. The methodology is used to propose strategies for maximising performance gain with quantitative results to support the conclusions.

The framework mixes hard systems modelling into a soft approach providing a method for managing complexity and facilitating learning points for the development of future wireless systems.

Dedication and acknowledgements

This work was supported by the University of Bristol, EPSRC funded Industrial Doctorate Centre in Systems (Grant EP/G037353/1) and Toshiba Research Europe Limited. In addition, I am indebted to the following people:

Professor Eddie Wilson for fantastic supervision, stimulating academic conversation and endless words of wisdom. Without your support, guidance and motivation this thesis would certainly not exist.

Professor Angela Doufexi for the academic guidance, support and feedback, always prompt and detailed.

Professor Mahesh Sooriyabandara for support and advice throughout every aspect of the project.

Colleagues at Toshiba and the University of Bristol for your hints, tips, recommendations and good conversation over coffee from which this thesis has benefited greatly.

Bryony for being the most wonderful and supportive partner.

My parents for your love and endless support throughout my life and academic career.

AUTHOR'S DECLARATION

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Research Degree Programmes and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

SIGNED: ..... DATE: .....

TABLE OF CONTENTS

Page

List of Tables **xi**

List of Figures **xiii**

List of Publications **xix**

List of Abbreviations **xxi**

1 Introduction **1**

1.1 Industrial Context and System Boundary 2
1.2 Hard and Soft Systems Engineering 3
1.3 Research Overview: A Systems Approach 4
1.4 Thesis Structure 6

2 Planning for the 5G Era **11**

2.1 Introduction 12
2.2 The Context of 5G Development 13
2.2.1 Understanding the Historical Context 13
2.2.2 Understanding Current Development Timeline to 5G 15
2.2.3 Understanding the Technical Drivers Context 15
2.3 Understanding New Business Opportunities for the 5G Era 19
2.4 Problem Structuring Approach: Strategy Development for the 5G Era 20
2.4.1 A Complex Problem Context 20
2.4.2 Modelling 5G Era Technology Development 22
2.4.3 Understanding the Stakeholders of the 5G Era 23
2.4.4 Problem Structuring Approach to the 5G Era 26
2.5 Implications of Problem Structuring 29
2.6 Developing Understanding of the 5G Era Business Landscape 30
2.6.1 The Historical Context: Understanding Challenges to Traditional Business Models 30
2.6.2 Understanding State-of-the-Art Business Models 31

2.6.3 Applying Problem Structuring-Systems Analysis of Business Models to Further Aid Understanding . . . 34
2.7 Reflections on 5G Era Strategy Development . . . 36
2.8 Conclusions and Scope for the Remainder of the Thesis . . . 38
2.A Appendix . . . 40
2.A.1 Summary of Business Models . . . 40
2.A.2 Decision on Research Investment . . . 43

3 A Comparison of Modelling & Simulation Methodologies 47
3.1 Background: Wireless CSMA/CA Networks . . . 48
3.1.1 Layman's Guide to the CSMA/CA MAC Protocol . . . 48
3.1.2 Managing Competing Demands in CSMA/CA Wireless Networks . . . 51
3.1.3 Applications of and Issues with the CSMA/CA Protocol . . . 54
3.2 Review of Modelling and Simulation of Wireless Networks . . . 60
3.2.1 Review of Analytical Modelling of Wireless Networks . . . 62
3.2.2 Review of Normative Simulation of Wireless Networks . . . 67
3.2.3 Modelling Wireless Network Topologies . . . 68
3.2.4 Modelling Random Backoff in Wireless CSMA/CA Networks . . . 72
3.2.5 Simulation Performance Metrics in Wireless CSMA/CA Networks . . . 75
3.3 Identifying the Opportunity for Innovation . . . 79
3.3.1 Conceptual Simulation Modelling . . . 80
3.3.2 Why Use Modelling and Simulation? . . . 81
3.3.3 Comparison of Methodologies . . . 83
3.4 Conclusion . . . 92

4 A Conceptual Simulation Modelling Methodology 95
4.1 Requirements . . . 96
4.1.1 Conducting a Successful Modelling / Simulation Study . . . 96
4.1.2 Verification, Validation and Usefulness . . . 98
4.2 Conceptual Simulation Model . . . 99
4.2.1 Conceptual Simulation Model Parameters . . . 101
4.2.2 Network Representation: Link-Centric Models . . . 103
4.3 An Ensemble Simulation Methodology . . . 103
4.4 Parametrisation . . . 107
4.4.1 Controlling Error Level In NS3 . . . 108
4.4.2 Fixed Parameters: *Transmit* and *Listen* . . . 112
4.4.3 Uniformly Distributed Parameters: *Backoff* and *Countdown* . . . 113
4.4.4 Validation Experiment 1: 2 and 3 Links . . . 117
4.4.5 Validation Experiment 2: Service and Latency Distribution . . . 121

4.4.6 Validation Experiment 3: 1-20 Links . . . 121
4.5 Conclusion . . . 123
4.A Appendix . . . 125
4.A.1 Meta Code . . . 125

5 MAC Layer Performance Evaluation with Full-Duplex Nodes 127
5.1 Background . . . 129
5.2 Study 1: Full-Duplex MAC in Wireless Mesh Networks . . . 131
5.2.1 Simplified Mesh Network Setup . . . 131
5.2.2 Operational Scenarios . . . 132
5.2.3 Simulation Methodology . . . 134
5.2.4 Results . . . 135
5.2.5 Analysis . . . 135
5.3 Study 2: A Performance Evaluation of the Impact of Full-Duplex . . . 137
5.3.1 Simple Network Setup . . . 137
5.3.2 Operational Scenarios . . . 138
5.3.3 Modelling the Bottleneck and Hidden Node Problems . . . 140
5.3.4 Simulation Methodology . . . 141
5.3.5 Results . . . 142
5.3.6 Analysis . . . 142
5.4 Discussion . . . 145
5.5 Conclusion . . . 147

6 A Pragmatic Approach to CCA Threshold and TP Adaptation 149
6.1 Introduction . . . 149
6.2 Randomly Generated Networks . . . 151
6.2.1 Randomly Generated Networks Results and Analysis . . . 155
6.2.2 Analysis of Matrix Characteristics . . . 160
6.3 Simple Topology Networks . . . 162
6.3.1 Simple Topology Operational Scenarios . . . 165
6.3.2 Simple Topology Simulation Methodology . . . 167
6.3.3 Simple Topology Results . . . 167
6.3.4 Simple Topology Analysis . . . 169
6.4 Discussion . . . 171
6.5 Conclusion . . . 174

7 Effects of Routing on the Capacity of Multi-Hop Networks 177
7.1 Background . . . 178
7.2 A Theoretical Framework for Computing the Capacity Region . . . 180

7.3 Single OD Pair Tutorial . . . 183
7.4 Single OD Simulation Study . . . 195
7.4.1 Simulating Multi-Hop Networks . . . 196
7.4.2 Simulation Results and Analysis . . . 197
7.5 Designed Networks With Two OD Pairs . . . 200
7.6 Random Networks With Two OD Pairs . . . 202
7.7 Discussion . . . 206
7.8 Conclusions . . . 207
7.A Appendix . . . 208
7.A.1 Beamforming Results . . . 208

8 Conclusions **209**
8.1 Summary of Research and Original Contributions . . . 209
8.2 Overall Context of Main Findings . . . 211
8.3 Opportunities for Further Research . . . 213
8.A Appendix . . . 218
8.A.1 Summary of Contributions from Each Chapter . . . 218

Bibliography **221**

List of Tables

Table **Page**

2.1 Summary of business models 42
2.2 Decisions on research investment strategy 46

3.1 CSMA/CA MAC protocol parameters from [29]. 54
3.2 Comparison of network modelling methods. 90

4.1 NS3 802.11b parameters [197]. 109
4.2 *Countdown* parameters required to replicate specified level of carrier sensing error (i.e., the probability an RTS generated does not lead to successful transmission of CTS, data packet and ACK) in CSMA/CA networks. 116
4.3 Calibrated conceptual simulation model parameters. 123

7.1 Network statistics for experimental ensembles of random wireless multi-hop networks. The parameter m describes the number of relay nodes. Other parameters are as defined in Sec. 7.2. Results reported to three significant figures; integer results are exact. 204
7.2 Capacity statistics for experimental ensembles of random wireless multi-hop networks, characterised by maximising $d_A + d_B$ subject to $d_A = d_B$ (GAME 2). Here the superscript u refers to the number used in the optimal assignment and the subscript AB refers to the number shared in the optimal assignment. The parameter $\tilde{n} \leq n$ is the number of non-redundant links. 204
7.3 Network statistics for experimental ensembles of random wireless multi-hop networks with beamforming capability only. The parameter m describes the number of relay nodes. Other parameters are as defined in Sec. 7.2. Results reported to three significant figures; integer results are exact. 208
7.4 Capacity statistics for experimental ensembles of random wireless multi-hop networks with beamforming capability only, characterised by maximising $d_A + d_B$ subject to $d_A = d_B$ (GAME 2). Here the superscript u refers to the number used in the optimal assignment and the subscript AB refers to the number shared in the optimal assignment. The parameter $\tilde{n} \leq n$ is the number of non-redundant links. 208

List of Figures

Figure **Page**

1.1 The inquiring / learning cycle of SSM adapted for this thesis. 5
1.2 The structure of this thesis 7

2.1 Historical timeline of the mobile telecoms industry highlighting some key moments in its development. 14
2.2 Timeline to 5G. As proposed by International Telecoms Union and 3rd-Generation Project Partnership. 15
2.3 Global mobile devices and connections growth forecast. 16
2.4 Global data traffic and WiFi offloading forecast. 19
2.5 Fifth generation era stakeholder influence versus interest diagram. 24
2.6 Hierarchical Process Model under development. 27
2.7 Hierarchical Process Model final. 28
2.8 A systemigram illustrating the link between the drivers for the 5G era and the development of business models. 35
2.9 The methodology of this chapter; from problem statement to decision making. 37

3.1 A successful CSMA/CA MAC protocol transmission along the link connecting A to B. 49
3.2 Flow diagram illustrating the relationship between the steps of the CSMA/CA MAC protocol, the conditions that determine if each step is successfully completed and the mechanisms by which each step may fail. 51
3.3 The backoff mechanism. (a) An illustration of how the backoff mechanism resolves contention for two links sharing the same channel. (b) An illustration of the backoff mechanism leading to an RTS collision. 52
3.4 Backoff retries sequence for 802.11. The figure illustrates the exponential growth in the backoff window. 53
3.5 The effect of SINR on the ability of two receivers to receive from their associated transmitters. 57
3.6 Hidden node problem. 59
3.7 Exposed node problem. 60

3.8 Node and link-centric models. A network of six nodes is connected by five unidirectional links. (a) A node-centric model / representation. (b) A link-centric model / representation. 71
3.9 An illustration of Laufer and Kleinrock's model progressing with time for three saturated links within carrier sense range. 73
3.10 Schematic of a typical capacity region for two demands. The diagram illustrates the performance of a system when demands are in contention for a limited resource. . . . 76
3.11 Laufer and Kleinrock two-link convex capacity region. 77
3.12 The range of research methodologies for wireless networks. 81

4.1 State diagram for each link of the conceptual simulation model. (a) Five-state and (b) four-state models. 100
4.2 Link-centric conceptual simulation model. (a) A six-node three-link topology. (b) A link collision matrix. (c) Illustrates the four-state model cycling through states as time progresses. 102
4.3 (a) Capacity region performance metrics for two links. (b) Plots proportion of a section through the demand space within capacity against demand. 104
4.4 (a) Illustrates discrete simulation points. (b) shows how the queue length changes as the simulation progress. 105
4.5 Simulation results demonstrating the need for a tolerance in the threshold used for judging if demand combinations fall inside or outside the capacity region. The threshold for judging if the queue of a simulation is growing over time is adjusted for the two simulations shown, with all other parameters kept the same. (a) Simulations are judged to be outside the capacity region if the mean gradient of the observed queue length over time is greater than zero. With this threshold (i.e., without any tolerance), only when the simulation ends with empty queues are they judged inside the capacity region. (b) Simulations are judged to be outside the capacity region if the mean gradient of the observed queue length over time is greater than 0.0001. 107
4.6 Carrier sensing error against propagation loss in NS3. 109
4.7 An illustration of why CSMA/CA can lead to a convex capacity region. (a) Shows a single saturated demand transmitting alone. (b) Shows two saturated demands sharing the channel without collisions. After successfully completing a transmission including RTS, CTS, data packet and ACK, a CSMA/CA transmitter must backoff for a period of time chosen from a uniform distribution, even when transmitting alone in a clear channel. If two transmitters are sharing the channel, the transmitter whose backoff counter decrements first will begin transmitting, hence the average delay between any two packets transmitting is reduced. The second transmitter's backoff is frozen for the duration of the transmission and resumes decrementing once it senses the channel clear again. (c) Illustrates the resulting capacity region. 110

4.8 An illustration of why introducing the *Listen* state enables a convex capacity region. (a) Shows the four-state model with a single saturated demand transmitting alone. (b) Shows the four-state model with two demands sharing the channel without collisions. (c) Shows the five-state model with a single saturated demand transmitting alone. Note, after each *Transmit* state the model must wait in *Listen* before it can *Transmit* again. (d) Shows the five-state model with two demands sharing the channel without collisions. Now the *Transmit* of the second link can begin while the first is in *Listen*, hence capacity is increased. (e) and (f) show the capacity regions for (b) and (d) respectively. 111

4.9 Validation of *Listen* parameter by comparison of conceptual simulation model with NS3. Capacity region plot for two competing links. 113

4.10 An illustration of why the *Countdown* state can cause collisions and how the *Backoff* state resolves collisions. Increasing the length of the *Countdown* will increase the chance of a collision. 114

4.11 An illustration of how increased level of carrier sensing error increases the concavity of the capacity region. The change in boundary reflects a change in the level of error. 115

4.12 Validation of the conceptual simulation model by comparison with NS3. Capacity region boundary plot for two competing links at three levels of carrier sensing error. . 117

4.13 Capacity region plots for three competing links. 118

4.14 Triangular sections of constant demand, identifying the proportion of the triangle within capacity. 119

4.15 Validation of the conceptual simulation model by comparison with NS3. Proportions of triangular section within the capacity region, as a function of total demand intensity, compared across three levels of carrier sensing error. 120

4.16 Each pair of plots shows on the left; the proportion of packets served; instantly (blue), in one to two packet lengths (red), in two to three packet lengths (black) and in three to four packet lengths (yellow). Comparing NS3 (markers) to the conceptual simulation model in Matlab (circles). On the right: an example latency distribution. 122

4.17 Validation for 1 to 20 competing links of equal demands by comparison of the conceptual simulation model with NS3. Plot shows the saturation points for 1 to 20 competing links of equal demands. 123

5.1 Simple three node mesh network for study 1. 131

5.2 Bidirectional and unidirectional full-duplex for the simple mesh topology. 132

5.3 Allowable (simultaneous) transmissions: (a) single transmission; (b) bidirectional full-duplex transmission; (c) unidirectional full-duplex transmission; (d) an extension of (c) with three nodes in full-duplex simultaneously. 133

5.4 Link collision matrices for each of the full-duplex simple mesh scenarios. 134

5.5 (a) Capacity region plot for Scenario 1 (half-duplex). (b) Triangular sections for constants demand intensity. (c) Corresponding plots of latency, showing how it diverges outside the capacity region. 136
5.6 Numerical results. (a) Proportion of triangular sections within the capacity region, as a function of total demand intensity, compared across the four Scenarios. (b) Table of the volume of the capacity region, and its extent along the line of equal demand to all transmitters. 136
5.7 Simple network setup for study 2. 137
5.8 Simultaneous transmissions occurring in full-duplex (a) Unidirectional full-duplex (b) Bidirectional full-duplex. 139
5.9 Link collision matrix. 140
5.10 Knowledge of the network link matrix. 141
5.11 (a) Capacity region plot for Bottleneck Problem. (b) A plot of demand against error rate. (c) Triangular sections for constant demand showing latency across the corresponding section for the two problems and three Scenarios investigated. 143
5.12 (a) Triangular sections for constant demand for Bottleneck simulation. (b) Proportions of triangular section within the capacity region, as a function of total demand intensity, compared across three scenarios for each of the two problems. (c) Table of volume of the capacity region. 144

6.1 Basic CCA threshold and TP adaptation. 150
6.2 Example of a randomly generated network. 152
6.3 Performance impact of applied modifications. 155
6.4 Impact of adaptation on throughput and fairness relative to legacy for 30 clients. . . . 156
6.5 Fairness impact of applied modifications. 157
6.6 Performance improvement / reduction range for the 80_{th} percentile, for all adaptation types. The potential improvement using the clustering coefficient is visible. 158
6.7 Mean power reduction for all networks for all adaptation types, normalised for the number of nodes. 159
6.8 An example graphical plot of 'knowledge of the network' and 'collision' matrices for 50 link topology. 160
6.9 The mean clustering coefficient and density of the collision matrices compared to mean throughput performance improvement for the top 10% of performing networks to the lowest 10% of performing networks in intervals of 10%. 161
6.10 Simple topology 1 with CCA threshold adaptation. 163
6.11 Simple topology 2 with CCA threshold adaptation. 163
6.12 Simple topology collision matrices. 165
6.13 Simple topology knowledge of the network matrices. 165
6.14 Capacity region plots for S1-S8. 168

6.15 (a) Proportions of triangular cross section within the capacity region, as a function of total demand intensity. (b) A plot of demand per client against failure rate for scenarios S1, S5 and S8. 169

7.1 An illustration of the maximal vector search method. 181
7.2 Basic multi-hop network. 186
7.3 TP multi-hop network. 187
7.4 TP multi-hop network (alternative variation). 188
7.5 Beamforming multi-hop network. 189
7.6 Full-duplex multi-hop network. 190
7.7 TP with full-duplex multi-hop network. 191
7.8 TP with full-duplex multi-hop network (alternative variation). 192
7.9 Full-duplex with beamforming multi-hop network. 193
7.10 Single OD multi-hop network capacity region simulation results. 198
7.11 Exemplar networks connecting two OD pairs. (a) Simplest network with one shared relay node and one route for each OD pair. (b) Extra relay nodes are added. (c) Eight interior nodes yielding multiple routes for both OD pairs. (d) Increase in capacity regions as relay nodes are added, assuming full-duplex and beamforming technology. 201
7.12 (a) Randomly generated network for two OD pairs and capacity regions when (b) two OD pairs route cooperatively and (c) demand is fixed equally across all routes for one OD pair whilst maximising the sum of flows for the other OD pair 203
7.13 Scatter plots for experimental ensembles of random wireless multi-hop networks. . . 205

8.1 A development Fig. 1.1 illustrating the cycle of inquiring and summarising the learning points generated in this thesis. 210
8.2 (a) Proportion of packets served instantly on arriving in the queue. (b) Log of latency distributions geometric decay rate against demand intensity. The intersection with the x-axis can indicate the capacity region boundary and corresponding demand intensity. 215
8.3 (a) Log of queue length distribution decay against demand intensity. (b) Proportion of time queue length is zero against demand intensity. Extrapolating either plot to the intersection with the x-axis can indicate the capacity region boundary. 215
8.4 Plot of log of mean inter service times and log of 1 / arrival rate against demand intensity. The point at which log of mean inter service time diverges from log of 1 / arrival rate plot indicates the boundary of the capacity region. 216

List of Publications

[1] W. Jones, M. Sooriyabandara, M. Yearworth, A. Doufexi, and R. E. Wilson, “Planning For 5G: A Problem Structuring Approach for Survival in the Telecoms Industry,” Syst. Eng., vol. 19, no. 4, pp. 301–321, Jul. 2016.

[2] W. Jones, R. E. Wilson, M. Sooriyabandara, and A. Doufexi, “Wireless Network MAC Layer Performance Evaluation with Full-Duplex Capable Nodes,” in Proceedings of the 12th ACM Symposium on QoS and Security for Wireless and Mobile Networks - Q2SWinet ’16, 2016, pp. 111–118.

[3] W. Jones and R. E. Wilson “Effects of Routing on the Capacity of Multi-Hop Wireless Networks,” In Proceedings of the 21st ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWIM ’18). ACM, New York, NY, USA, 155-162.

[4] W. Jones, R. E. Wilson, A. Doufexi, and M. Sooriyabandara “A Pragmatic Approach to Clear Channel Assessment Threshold Adaptation and Transmission Power Control For Performance Gain in CSMA/CA WLANs” IEEE Transactions on Mobile Computing, Jan. 2019.

List of Abbreviations

3G Third Generation Network

3GPP Third Generation Project Partnership

4G Fourth Generation Network

5G Fifth Generation Network

ACK Acknowledgement

ACM Association for Computer Machinery

AMPS Advanced Mobile Phone System

AODV Ad-Hoc On-Demand Distance Vector

AP Access Point

CCA Clear Channel Assessment

CDMA Code Division Multiple Access

CPU Central Processing Unit

CSMA Carrier Sense Multiple Access

CSMA/CA Carrier Sense Multiple Access with Collision Avoidance

CTK Contribution to Knowledge

CTS Clear to Send

D2D Device to Device

DAG Directed Acyclic Graph

DCF Distribute Co-ordinate Function

DES Discrete Event Simulation

DIFS DCF Interframe Space

DL Down-Link

DSC Dynamic Sensitivity Control

DSDV Destination-Sequenced Distance Vector

FDMA Frequency Division Multiple Access

FD-MAC Full-Duplex Medium Access Control

GCD Generic Constitutive Definition

GPRS General Packet Reconstruction System

GSM Global System for Mobile Communication

HD-MAC Half-Duplex Medium Access Control

IEEE Institute of Electrical and Electronic Engineers

IoT Internet of Things

IP Intellectual Property

ISM Industrial, Scientific and Medical

IT Information Technology

ITU International Telecommunications Union

HetNet Heterogeneous Network

HPM Hierarchical Process Modelling

LTE Long-Term Evolution

MAC Medium Access Control

MIMO Multiple-Input Multiple-Output

MMC Massive Machine Communication

NFV Network Function Virtualization

M2M Machine to Machine

MCS Modulation Coding Scheme

mmWave Millimetre Wave

MTC Machine Type Communication

MU-MIMO Multiple-User Multiple-Input Multiple-Output

NES Nash Extension Solution

NGMN Next Generation Mobile Network

NS2 Network Simulator 2

NS3 Network Simulator 3

OD Origin-Destination

OFDMA Orthogonal Frequency Division Multiple Access

OFDM Orthogonal Frequency Division Multiplexing

OLSR Optimised Link State Routing

OPNET Optimised Network Engineering Tools

OTT Over the Top

PAN Personal Area Network

PAS Purposeful Activity System

PHY Physical Layer

PPP Poisson Point Process

PSM Problem Structuring Method

QoS Quality of Service

RAT Radio Access Technology

R&D Research and Development

RSSI Received Signal Strength Indication

RTS Request to Send

SDN Software-Defined Network

SIFS Short Interframe Space

SINR Signal to Interference Plus Noise Ratio

SSM Soft Systems Methodology

SUMO Simulation of Urban Mobility

TDMA Time Division Multiple Access

TP Transmission Power

TRL Technology Readiness Level

UAV Unmanned Air Vehicles

UL Up-Link

UMTS Universal Mobile Telecommunications System

VANET Vehicular Ad-Hoc Network

V2V Vehicle to Vehicle

V2I Vehicle to Infrastructure

V2X Vehicle to Anything

VNI Visual Networking Index

VR Virtual Reality

WiFi Wireless Fidelity

WLAN Wireless Local Area Network

WSN Wireless Sensor Network

CHAPTER 1

INTRODUCTION

Systems engineering is the field that focuses on how to design and manage complex systems [181, 188]. A system consists of a number of parts or items that are grouped together, interrelated and interdependent; it may be natural or man-made. Characteristically, a system is often more than the sum of its parts. That is, when brought together in a system, the parts are able to produce emergent behaviours that could not be achieved otherwise [230]. Due to the interconnected nature, implementing a change in one part of the system usually impacts on other parts, and potentially the whole system.

For systems engineers, the objective is to discover a system's underlying principles by studying its dynamics, constraints, conditions, etc [34, 138]. This applies when considering either hard (i.e., physical / technical) or soft (i.e., management / social) systems. Often the objective is that the system can then be modified so as to operate in an optimal manner.

Systems modelling is a basic principle in systems engineering [18, 71, 177] and has similarities with modelling methods common in mathematics and other forms of engineering. When developing a model, the system is the entities under concern and inclusion to or exclusion from the system is at the modeller's discretion and dependent on their intentions. The aim, typically, is to make the system, or parts of it, easier to understand, quantify, visualise and predict. This type of model typically involves an abstraction of the real world, referenced to existing and commonly accepted knowledge. No model of a system will capture all features of the real system of interest, and no model of a system needs to include all entities belonging to a real system.

This thesis applies the tools and methods of systems engineering to study wireless communication systems. Systems modelling approaches are used to explore complex aspects of wireless communications systems and subsequently, recommendations are made for improvements.

1.1 Industrial Context and System Boundary

This engineering doctorate is supported by the University of Bristol, EPSRC funded Industrial Doctorate Centre in Systems (Grant EP/G037353/1) and Toshiba Research Europe Limited.

Toshiba Research Europe Limited is an industry research laboratory, a subsidiary of Japanese multinational conglomerate Toshiba whose headquarters are in Tokyo, Japan. Toshiba's diversified products and services include industrial and social infrastructure systems, lighting, logistics, electronic components, consumer electronics, information technology and communications equipment and systems, power systems, household appliances, office equipment and medical equipment. This engineering doctorate was conducted in association with and based in the organisation's Bristol Research Laboratory. The interest of the laboratory is wireless communications research and development. The objective of the laboratory is to develop technologies to incorporate into future products that can be sold and to build an intellectual property portfolio for potential licensing.

From 2016 to 2018 an 8% compound annual growth rate in the number of wireless connected devices has been observed. It is estimated that by the year 2021 there will be 11.6 billion wireless connected devices worldwide, exceeding the projected population of the world (7.8 billion) [21]. This presents a significant market opportunity for technology-focused organisations such as Toshiba.

From the organisation's perspective, research is pursued to develop new process and product innovation, and further to develop broader capabilities to assimilate and exploit externally available information, i.e., to learn [61]. Organisations need to create innovations in order to meet today's market demands and further to ensure their long-term survival [258].

A narrow perspective sees wireless communications systems as purely technical systems. This draws a tight system boundary around the physical infrastructure that facilitates communications. An alternative perspective draws a wider system boundary and views wireless communication systems as complex socio-technical systems [270]. They can be viewed as a physical system embedded within a surrounding social and economic system.

From the perspective of Toshiba Research Europe Limited, the growing wireless communications industry presents an opportunity for innovations that could provide commercial returns. Although the lab's output focuses on the technical system and its operation, it is necessary to consider the influence of external factors on the wider social and economic system.

This thesis explores socio-economic aspects of the wireless communications industry so as to identify key areas that are likely to provide a return on research investment for Toshiba. The thesis then provides a detailed technical investigation into some of those key areas.

1.2 Hard and Soft Systems Engineering

When referring to hard systems, the word 'system' is used to label something taken to exist in the world outside ourselves [50]. The inherent assumption is that the world consists of a set of such interacting systems, some of which can potentially be engineered to work better.

Conversely, when considering soft systems, the inherent assumptions are quite different. Now, the world is taken to be unclear, very complex and incomprehensible in full. However, a process of inquiry into it, and a method of coping with it can be designed as a learning system. Hence in this situation, the word 'system' is no longer directly applied to the world but instead to the methodology of dealing with the world [48, 50].

This shift of complex, dynamic behaviour exhibited by systems from the world to the process of inquiry into the world is the crucial intellectual distinction between the two fundamental forms of thinking associated with systems engineering (hard and soft). Peter Checkland [48] explained that from the perspective of system engineering, when thinking of a hard system, the starting point is a structured problem and the objectives of the system are clear and stable. On the other hand, when thinking about a soft system, the starting point is unstructured and the objective is to bring structure, facilitating debate about actions in problem-solving within the system.

Hard and soft approaches differ in the view they take of models themselves. In a hard approach, the typical assumption is that the model is a comprehensive representation of part of the real world, accepting however that it will be a simplified abstraction. This approach has resonance with methods common in mathematics and various fields of engineering where there is interest in developing models to obtain quantitative results. By contrast, in soft approaches, the idea is that models are developed so as to facilitate debate about possible action. Hence, the main concern is that the models should be useful.

Soft approaches do not necessarily guarantee a set of recommendations or a particular product, instead, they stress the importance of learning. Their focus is on providing a tool for coping with complexity and learning so that performance can be improved. Tangible findings or products may emerge from this type of approach, or further to the learning provided. Hence soft methods are typically considered as cyclical and part of a continual progression.

Multi-methodological approaches [177–179, 217, 218], mixing hard systems modelling into a soft approach, are well understood and entirely feasible [146]. The approaches can feed into each other at different times and may enrich each other.

This thesis mixes hard systems modelling into a soft approach. While significant effort is spent developing an accurate representation of the real world, the focus is on developing a research method that facilitates learning. This approach is applied in the context of wireless communication systems, such that recommendations can be provided to facilitate their development.

1.3 Research Overview: A Systems Approach

Soft systems methodology (SSM) is an example of a methodology used to support and to structure thinking about, and intervention in, complex problems. It was developed by Peter Checkland [49] providing an approach which facilitates decision making in relation to complex systems. As Ison comments [121], "SSM is not a tool or technique to be used occasionally but a way to think and act every day". Fig. 1.1 adapts Checkland's learning inquiry model [50], illustrating the structured approach this thesis presents for addressing the complex challenges faced by a research and development organisation in the wireless industry, exploring the next generation of communications. Further, the principles of Checkland's model [50] are adapted to illustrate the approach of this thesis as follows:

- The real world (i.e., the wireless communications industry) is a complex system; consisting of relationships between many sub-systems.
- Models of purposeful activity based on explicit worldviews (i.e., a perception of what is going on in the real-world) can be used to explore complex systems.
- An inquiry is structured by questioning a perceived situation using the models as a source of questions.
- Actions to improve are based on finding acceptable accommodations (i.e., changes that can be made to the system, that despite in some cases having a slight negative impact, in general, have a positive impact overall).
- Inquiry in principle is never-ending and best conducted with a wide range of interested parties. Modelling facilitates engagement with key stakeholders.

Research and development organisations are faced with the challenging question of how to develop a strategy for research investment in the complex development space of fifth generation (5G) era technologies. This thesis begins by exploring this question and proposes a methodology for doing so. The approach used is based on the use of Problem Structuring Methods and focuses on developing a clearer understanding of the development landscape of the telecoms industry for informed decision making. The approach is used to identify key areas of wireless research for the 5G era.

Further, this thesis proposes a novel conceptual simulation modelling methodology, designed to aid researchers in the environment of a corporate research and development lab exploring next-generation wireless technologies with the goal of developing innovations and intellectual property that can bring commercial success. The thesis explains why this is a useful approach for exploring proposed key technologies of the 5G era in the context of Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) wireless networks. Significant effort is required to design, build, and validate a conceptual simulation modelling methodology before it can be used.

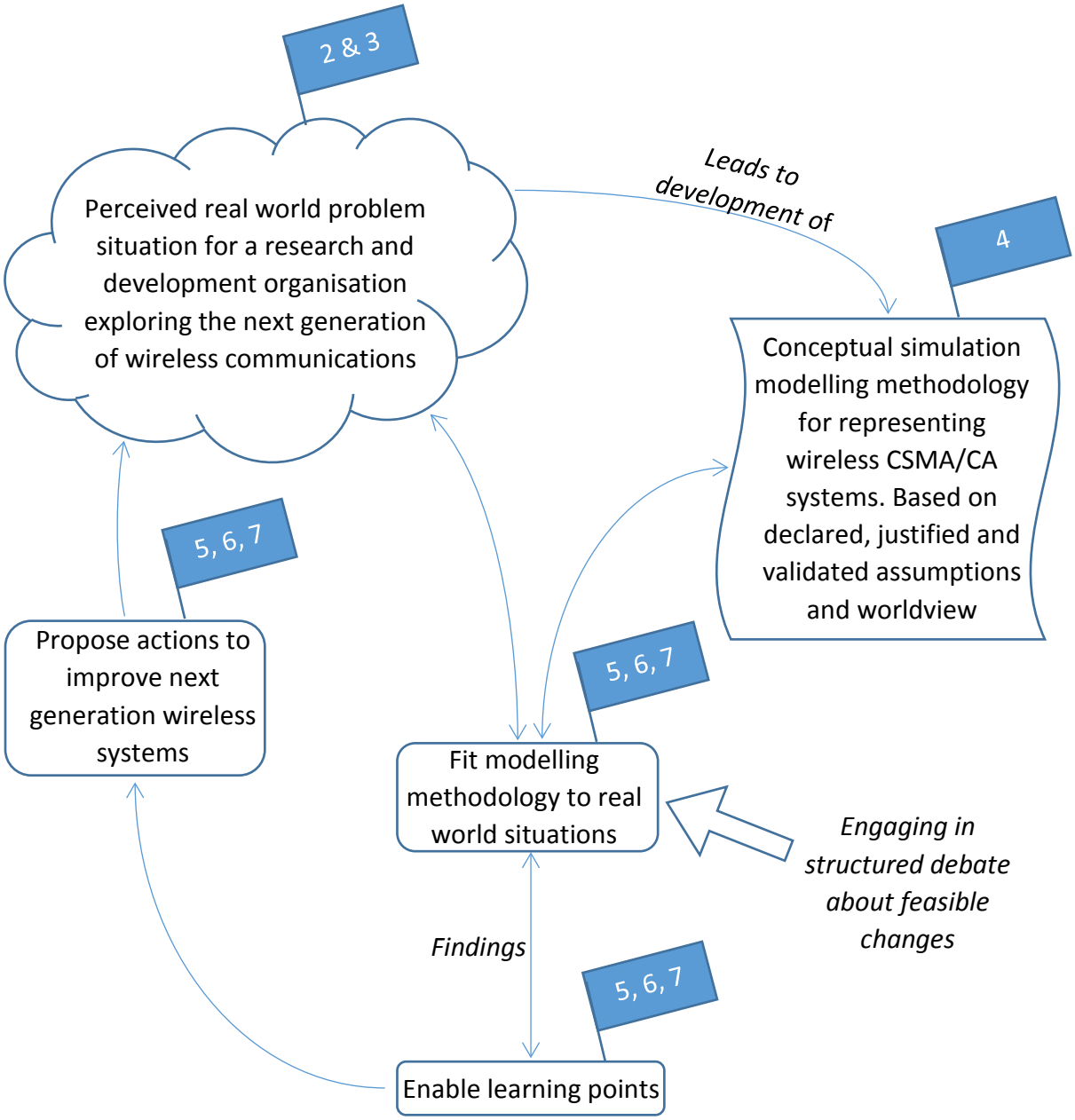


FIGURE 1.1. The inquiring / learning cycle of SSM adapted for this thesis. Adapted for this thesis from: "P. Checkland, Systems Thinking, Systems Practice (30 Year Retrospective). Oxford UK: Wiley, 1999." The blue flags, attached to each stage of the model, indicate the chapters that each stage is addressed in.

The usefulness of the conceptual simulation methodology, in enabling learning points, is demonstrated through its application to three identified key areas of wireless technology research, namely (1) full-duplex, (2) clear channel assessment and transmission power adaptation and (3) routing in multi-hop wireless mesh networks. Whilst each technology is capable in principle of improving system performance, often the gain is limited when implementing in a mesh network. The methodology is used to propose strategies for maximising performance gain with quantitative results to support conclusions.

1.4 Thesis Structure

Following this introductory chapter, the remainder of the thesis is structured as follows; Chapters 2, 3 and 4 should be read sequentially ahead of Chapters 5, 6 and 7 which can be considered in parallel. Chapter 8 concludes the thesis. Fig. 1.2 provides an overview of the thesis structure highlighting the key areas of interest for each chapter in terms of both the development of the modelling and simulation approach and further the findings in relation to wireless networks. A summary of each chapter is as follows.

Chapter 2: *Based on: W. Jones, M. Sooriyabandara, M. Yearworth, A. Doufexi, and R.E. Wilson, "Planning For 5G: A Problem Structuring Approach for Survival in the Telecoms Industry," Syst. Eng., vol. 14, no. 3, pp. 305-326, Jul. 2016.*

This chapter examines the application of systemic problem structuring methods for the development of a research strategy in response to the challenges of 5G era communication networks. The chapter proposes a methodology for strategic decision making to help corporate research and development organisations identify research areas likely to provide a return on investment. The key stakeholders, objectives, technologies, and boundaries from existing literature are identified and problem structuring based on hierarchical process modelling is used to explore the dependency of certain features of 5G era networks on specific technologies, giving an indication of the importance of certain technologies over others and thus insight into where to place research effort. The hard technical challenges of the 5G era are discussed and equally the importance of the soft social and business challenges explored. For context, how the 5G era will provide a platform for innovations is explained and how new and existing businesses may use this to their advantage is discussed. Problem structuring is used to explore how the challenges and opportunities of future wireless systems are related to the process of developing new business models.

Chapter 3: This chapter defines and proposes a conceptual simulation methodology and argues why it is a useful approach for research in the development of CSMA/CA wireless networks. Some background on CSMA/CA networks is provided and further extensive literature review of modelling and simulation research in relation to them is presented. Following

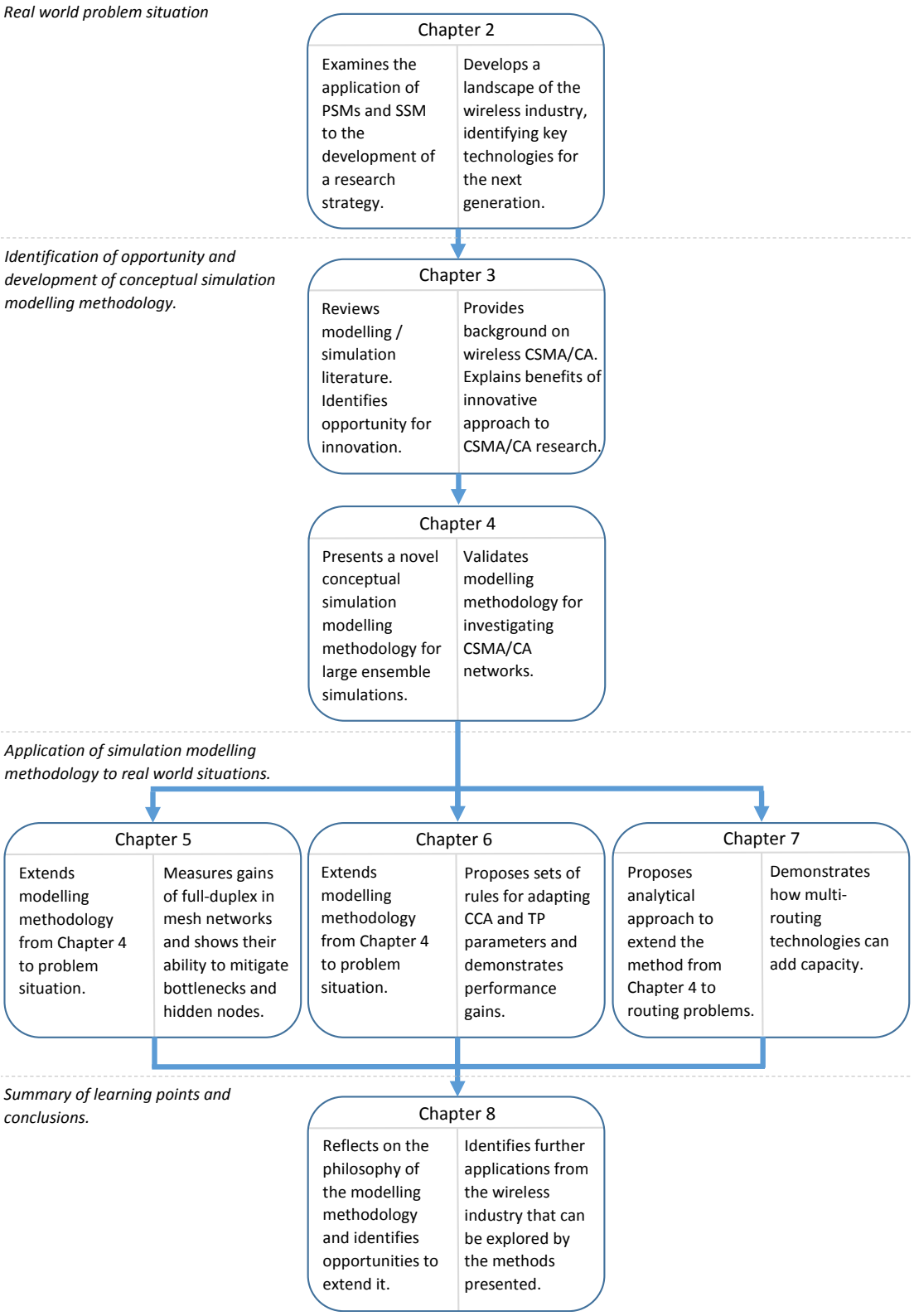


FIGURE 1.2. The structure of this thesis. For each chapter, the bubble explains the interest from a modelling and simulation perspective (on the left) and a wireless communication perspective (on the right).

examination of the existing literature, categories of research methodology, namely analytical modelling and normative simulation modelling are compared with the proposed conceptual simulation modelling approach. The inherent assumptions of these methodologies are discussed, in parallel with their suitability and potential contribution to the development of next-generation wireless networks. The methodologies' capabilities are discussed. It is identified that a validated conceptual simulation modelling methodology uses the strengths of problem structuring methods to overcome limitations of purely analytical techniques.

Chapter 4: A conceptual simulation modelling methodology for representing wireless CSMA/CA networks is presented in detail. The methodology is designed to aid researchers in the environment of a corporate research and development lab, exploring next-generation wireless technologies, with the goal of developing innovations and intellectual property that can bring commercial success. The conceptual simulation modelling methodology is shown to produce results comparable to those in existing literature for specific test cases. Further quantitative validation is given by comparison with normative simulation results generated using Network Simulator 3.

Chapter 5: *Based on: W. Jones, R.E. Wilson, M. Sooriyabandara, and A. Doufexi, "Wireless Network MAC Layer Performance Evaluation with Full-Duplex Capable Nodes," in Proceedings of the 12th ACM Symposium on QoS and Security for Wireless and Mobile Networks - Q2SWinet'16, 2016, pp. 111-118.*

Two exemplar studies demonstrate the potential impact on network performance that full-duplex can provide. The first demonstrates the potential capacity gain that full-duplex can provide applied in various forms to a highly simplified mesh network set-up, showing that full-duplex alone can increase the capacity of a network; however, when combined with appropriate interference management, the gain is much more significant. Existing work speculates at this result but this study is the first to quantify it. The majority of the capacity gain was shown to occur at asymmetric demand combinations which would not have been observed only considering equal or saturated demands. The second study focuses on common issues in current communications networks; bottlenecks and hidden nodes, demonstrating the potential performance gains that full-duplex nodes can offer via a simple example of two clients and an access point. It is shown that introducing full-duplex access points alone mitigates against the problem of bottlenecks, reduces the impact of hidden nodes and can increase the capacity of a network. When full-duplex access points are able to work with full-duplex clients, the capacity gain is much more significant; however, it is shown that much of this capacity gain occurs at uneven demand combinations. When the demand to all nodes is equally high, the introduction of full-duplex capability to clients is shown to increase the number of transmission attempts resulting in a significantly increased number of collisions and reduced network performance. Further,

it is observed that at low traffic levels, a full-duplex access point may improve goodput by simply transmitting a busy tone to silence other transmissions whilst it receives, mitigating against the hidden node problem.

Chapter 6: *Based on: W. Jones, R. E. Wilson, A. Doufexi, and M. Sooriyabandara "A Pragmatic Approach to Clear Channel Assessment Threshold Adaptation and Transmission Power Control For Performance Gain in CSMA/CA WLANs" IEEE Transactions on Mobile Computing, Jan. 2019.*

A practical set of rules for adapting clear channel assessment and transmit power parameters are proposed and evaluated by simulating ensembles of randomly generated wireless networks and collecting throughput statistics. The rules' performances depend strongly on network topology, with increases in throughput in many cases. However, networks with a high clustering coefficient are often adversely affected by the adaptations. But simulations of small-scale networks show that apparently adverse adaptations may still yield benefits for uneven demand combinations. Finally, it is shown that throughput is not usually correlated with the number of hidden or exposed nodes in any non-trivial network set-up.

Chapter 7: *Based on: W. Jones and R. E. Wilson "Effects of Routing on the Capacity of Multi-Hop Wireless Networks," In Proceedings of the 21st ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWIM '18). ACM, New York, NY, USA, 155-162.*

This chapter considers multi-hop wireless mesh networks and examines whether capacity may be improved by distributing the data flows for each origin-destination (OD) pair across multiple routes. The network geometry together with rules for transmit power (and thus range) and the application of technologies such as beamforming or full-duplex are used to derive a conflict matrix that describes which pairs of transmissions (or links) are compatible, in that they can occur simultaneously without collisions or other conflicts. The paper then presents a theoretical framework for computing the capacity region. Firstly, the conflict matrix is used to derive the maximal sets of compatible links, and secondly, these are used to derive a system of linear inequalities that bound links' data flows. These steps are computationally expensive, but it is shown how the systems complexity collapses when re-expressed in terms of flows on routes. The theory is explained in terms of a simple 'Braess' network with a single OD pair, and compared with simulation results. The study then shows larger-scale numerical examples with two OD pairs whose data flows have some nodes in common and thus contend with each other. In some cases, extra capacity can be gained if the OD pairs distribute their traffic over several routes. The possibilities are examined via a set of linear programs that model (i) cooperative behaviour; (ii) the optimisation of one OD pair when presented with a fixed route assignment by the other;

and (iii) variants of these games when both OD pairs are in contention with background (single-hop) traffic.

Chapter 8: This chapter concludes the thesis. The contributions to knowledge (CTK) are summarised and possible future work is discussed.

CHAPTER 2

Planning for the 5G Era: A Problem Structuring Approach for Survival in the Telecoms Industry

This chapter is based on: W. Jones, M. Sooriyabandara, M. Yearworth, A. Doufexi, and R.E. Wilson, "Planning For 5G: A Problem Structuring Approach for Survival in the Telecoms Industry," Syst. Eng., vol. 14, no. 3, pp. 305-326, Jul. 2016.

This chapter examines the application of systematic problem structuring methods for the development of a research strategy in response to the challenges of next-generation communication networks. A methodology for strategic decision making to help corporate research and development (R&D) organisations identify research areas likely to provide a return on investment is proposed. The key stakeholders, objectives, technologies, and boundaries from existing literature are identified and problem structuring based on hierarchical process modelling [66] is used to explore the dependency of certain features of fifth generation (5G) [10] era networks on specific technologies, giving an indication of the importance of certain technologies over others and thus insight into where to place research effort. The hard technical challenges of 5G era technologies are discussed and equally the importance of the soft social and business challenges explored. For context, how the 5G era will provide a platform for innovations is explained and how new and existing businesses may use this to their advantage is discussed. Problem structuring [4, 179, 216, 269] is used to explore how the challenges and opportunities of future wireless systems are related to the process of developing new business models.

2.1 Introduction

The telecommunication industry has been growing at an unprecedented rate over the last three decades, during which the world has seen the development of four generations of cellular technologies. While the industry is continuing the deployment of fourth generation systems, global research and development efforts (i.e., a number of projects in Europe, United States, and Asia) are well underway to develop fifth generation (5G) telecommunication systems [84]. Furthermore, the International Telecoms Union (ITU) has also started a standardisation effort which works toward specifying the overall framework and objectives of the future development of the telecommunications network, discussing aspects such as technology trends and the technical feasibility of new frequency bands.

It is acknowledged widely that the underlying difference between the 5G era and previous eras of wireless technology is that it will consist of more than just one radio access technology (RAT). The trend of using multiple RATs for providing network services is already well established and practised widely by operators. For instance, many operators own WiFi hotspot services in addition to their cellular networks. This architecture is well supported by the wide availability of smart mobile devices equipped with multiple radios such as cellular (3G/4G), WiFi, and Bluetooth. This observation is complemented by the description of the technology as highly integrative between the new 5G air interface with Long-Term Evolution (LTE) and WiFi [10]. It is suggested that 5G era wireless communication systems will require a mix of new system concepts to boost their spectral and energy efficiency [116].

In line with the above observations and predictions, this work hypothesises that the 5G era communication system will be far more complex than any of its predecessors, which mostly use homogeneous technologies with vertically integrated market structures. In contrast, 5G era systems will be complex socioeconomic-technical systems which will integrate different communications subsystems that use one or few specific technology solutions and network architectures [164], facilitating multiple service operators and a range of applications to coexist, enabling innovative market opportunities.

This chapter explores the question R&D organisations are faced with of how to develop a strategy for research investment in the complex development space of 5G era technologies and proposes a methodology for doing this. The approach used is based on the use of Problem Structuring Methods (PSMs) and focuses on developing a clearer understanding of the development landscape of the telecoms industry for informed decision making. The work takes the perspective that the historical development of the industry will also provide some insight into the shape of its future. The influence of key state-of-the-art technologies and major projects is considered.

The chapter is organised as follows. In Sec. 2.2, a landscape of the telecoms industry is developed by understanding its historical development, examining the proposed development towards the 5G era and reviewing the industry state-of-the-art. In Sec. 2.4, problem structuring and modelling techniques, and their application to the problem in hand are discussed. Furthermore,

stakeholder mapping and Hierarchical Process Modelling (HPM) are applied. In Sec. 2.5, the implications of the problem structuring are discussed. In Sec. 2.6, reflections on 5G era strategy development are provided. In Sec. 2.7, the findings and methodology are discussed and in Sec. 2.8, conclusions are given.

2.2 The Context of 5G Development

2.2.1 Understanding the Historical Context

In trying to understand how the telecoms industry may progress, it is sensible to look at its past evolution. Reviewing the history provides context and helps to develop a conceptual landscape of the telecoms industry (see Fig. 2.1).

The technology of the telecoms industry is always chasing rising data demand. Data demand has increased faster than technology can keep up [10, 21]. Development of new technologies seems to be accelerating. The time from a standard being formed to a technology being realised has decreased over the lifetime of the industry. This is presumably due to the increased size of the industry and the associated economy attracting an increased number of participating research and development organisations. Notably, the rate of development of physical layer (PHY) technologies has slowed in the latter period of this timeline, primarily due to link-level throughputs (i.e., rate of successful message delivery measured in bits per second (bit/s or bps) or data packets per second (p/s or pps)) approaching Shannon's limit [209], whereas improvements to medium access control (MAC) layer technologies have accelerated. Heterogeneity has been a developing trend throughout the telecoms timeline and so the assumption that this trend will continue into the 5G era seems sensible [10].

Innovation has been key to success in the telecoms industry. Staying ahead of the technology curve [14] and reaching the market early with new technologies have seen a benefit in attracting new customers. Ambidextrous organisations [200, 201], such as Samsung, able to explore future options, build resources, and manage ongoing R&D, while growing the profitability of the ongoing business, have become significant in the industry. Innovation alone has not been sufficient to stay at the top. Former industry leader Nortel Networks R&D labs produced the world's first entirely digital phone network and dominated the fibre-optic network in the early 2000s, but due to lack of a clear strategy, descended into mediocrity [235]. In 2009, they filed for protection from creditors under bankruptcy. Motorola, a key innovator in the early years of telecoms, having stagnated in growth and witnessed a decline in market share, was bought out by Google [99, 184]. Google was interested in acquiring the company's patent portfolio and increasing the popularity of their Android operating system [156]; they proceeded to sell-on the remaining aspects of the company.

Large organisations within the telecoms industry have often portrayed a monopoly mentality looking to dominate the industry by growing their assets. However, in recent years this strategy

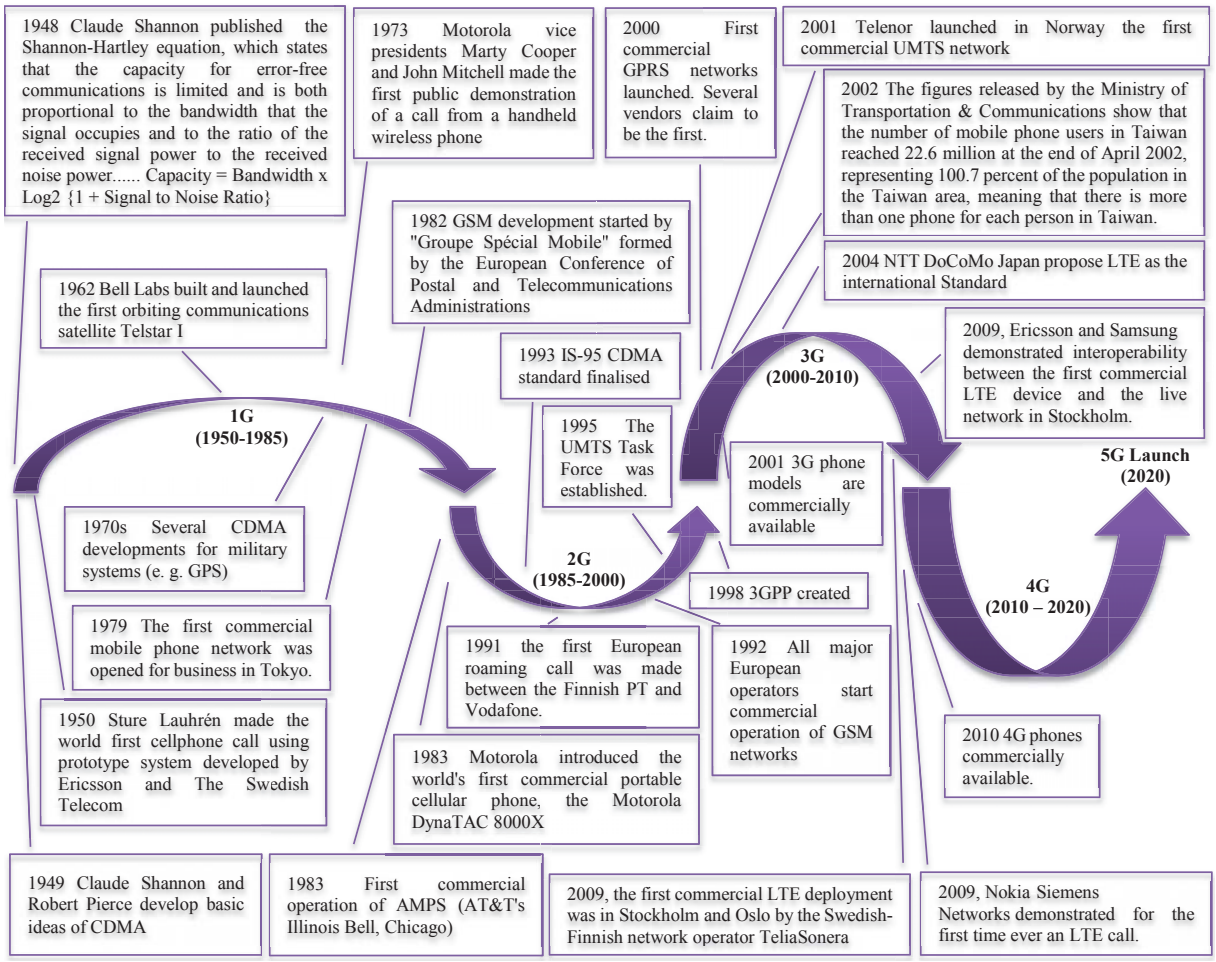


FIGURE 2.1. Historical timeline of the mobile telecoms industry highlighting some key moments in its development. Key terms; CDMA - Code Division Multiple Access; AMPS - Advanced Mobile Phone System; GPRS - General Packet Reconstruction System; UMTS - Universal Mobile Telecommunications System; 3GPP - Third-Generation Project Partnership, GSM - Global System for Mobile Communications; LTE - Long-Term Evolution. The information for the timeline was adapted from UMTS World [256] and Wireless History Foundation [271].

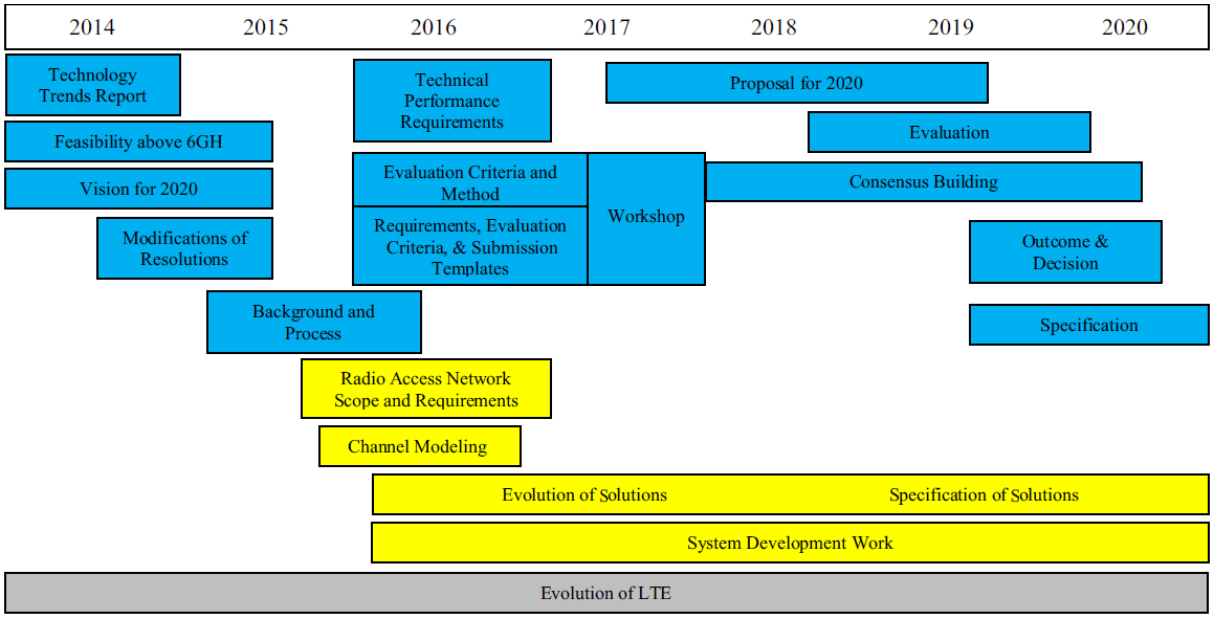


FIGURE 2.2. Timeline to 5G. As proposed by International Telecoms Union (ITU) in blue and 3rd-Generation Project Partnership (3GPP) in yellow.

has been disrupted by the emergence of over-the-top (OTT) providers which have in many cases become more profitable [163].

2.2.2 Understanding Current Development Timeline to 5G

Having briefly reviewed the history, it is now important to consider the current state-of-the-art in the Telecoms industry.

Fig. 2.2 shows a timeline to development of the 5G cellular network as followed by ITU (in blue) and 3rd Generation Project Partnership (3GPP) (in yellow). These two regulatory bodies are heavily influenced by many significant Engineering and Physical Sciences Research Council (EPSRC) and international projects already underway; these include TOUCAN [251], 5G-Xhaul [2], 5G Now [1], Fed4FIRE [83], NDFIS [190], FIBRE [85], FUTEBOL [90], STRAUSS [185], GEANT [96] and The METIS 2020 Project [174], and many others. These projects are focusing primarily on the architecture of 5G systems. Some of the technology drivers and business opportunities discussed in these projects are expanded upon in the following sections.

2.2.3 Understanding the Technical Drivers Context

The 'Mobile and Wireless Communications Enablers for the Twenty-Twenty Information Society' (METIS) project identifies some technical requirements envisaged for 5G era networks [202]. These are:

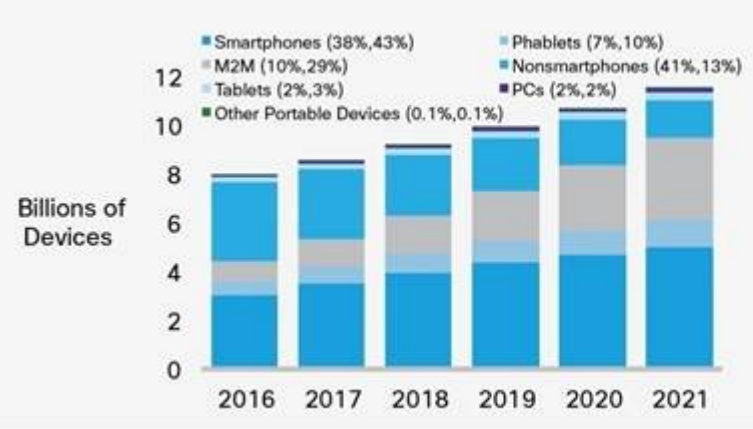


FIGURE 2.3. Global mobile devices and connections growth forecast. Figure reproduced from: T. J. Barnett, A. Sumits, S. Jain, and U. Andra, "Cisco Visual Networking Index (VNI) Update Global Mobile Data Traffic Forecast," Vni, pp. 2015–2020, 2015 [21].

- 1000 times higher mobile data volume per area.
- 10–100 times higher number of connected devices.
- 10–100 times higher user data rate.
- 10 times longer battery life for low-power massive machine communication (MMC).
- 5 times reduced end-to-end latency [202].

As cellular and WiFi have evolved, the two have become increasingly interwoven in modern telecoms systems and will continue to do so for the 5G era. This evolution of wireless technologies is driven by the need to accommodate the significant and continually increasing number of connected devices (see Fig. 3.5), and the ever increased demand for data (see Fig. 3.6). Further drivers include a demand for higher speeds, reduced latency, improved (total) coverage, improved quality of service (QoS), and continued service while in high-density population areas or while travelling.

The proposed Institute of Electrical and Electronics Engineers (IEEE) standard 802.11ax (expected 2019) is set to replace the existing 802.11n (released 2009) and 802.11ac (released 2013) standards as the wireless local area network (WLAN) standard reflecting the increased performance required [141] to satisfy anticipated user characteristics and meet the forecast data traffic. By examining a number of use cases and proposed scenarios, four key requirements for the IEEE 802.11ax amendment have been identified as follows [23, 141]:

- Coexistence

WLANs operate as unlicensed devices in the industrial, scientific, and medical (ISM) band. The ability to coexist with other wireless networks that operate in the ISM band is essential. Wireless sensor networks (WSNs) and personal area networks (PANs) are already becoming ubiquitous here and plans have been proposed by operators to extend LTE deployment into the ISM band [3]. The coexistence of WLANs presents a significant challenge that any IEEE amendment must address by attempting to minimise the interference they generate and be capable of operating in a high interference environment.

- Higher Throughput

  Where previous amendments focused on improving aggregate throughput, the 802.11ax amendment focuses on improving metrics that reflect user experience. Improvements will be made to support use cases such as wireless offices, outdoor hotspots, dense residential apartments, and stadiums [72]. These high-density environments require improved efficiency in using the finite PHY resources to enable higher throughput per user. New wireless technologies such as orthogonal frequency-division multiple access (OFDMA) [157, 205], beamforming and sectorisation [102, 267] [82], dynamic sensitivity control (DSC) [43], and advanced multiple antenna techniques such as multi-user multiple-input and multiple-output (mu-MIMO) [262] along with millimetre wave (mmWave) [210] are proposed to maximise the available capacity (i.e., the maximum amount of data that may be transferred between between two nodes).

- Energy Efficiency

  As with all prior amendments it is required that there be an improvement in power efficiency per data rate, however, these are becoming more difficult to sustain as data rates have rocketed and will continue to do so for 802.11ax. The target in IEEE 802.11ax-2019 is, at least, to match the energy consumption upper limit of previous amendments and not exceed it. This requires both new low-power hardware architectures [207] and new low-power PHY/MAC functionality.

  Improvements in efficiency may be found in the carrier-sense multiple access with collision avoidance (CSMA/CA) channel access protocol: the backoff process, packet headers, inter-frame spaces, collisions, and retransmissions process all present opportunities for savings. In comparison to the current MAC protocol, it may be possible to significantly increase the effective time that a node spends transmitting data every time it accesses the channel and so reduce energy consumption. IEEE 802.11ax-2019 aims to include several solutions to mitigate against MAC layer overheads. These include packet headers, aggregation, and piggy-backing [141], simultaneous transmit and receive (full-duplex) [207], [124] and further transmission power (TP) control and clear channel assessment (CCA) threshold adaptation [56, 248, 282].

- Backward Compatibility

  A requirement of any WLAN implementing IEEE 802.11ax is that it must also support devices using any previous IEEE 802.11 PHY/MAC amendments. Thus mechanisms need to be implemented allowing for backward compatibility (i.e., common frame headers and transmission rates). This is a source of inefficiency.

Key drivers and enabling technologies of the 5G era, for cellular and WiFi systems, have been identified at a PHY and a MAC layer [10, 23, 84, 97, 116, 141, 202]. Drivers include the significant continuing increased data demand and increasing number of connected devices (the Internet of Things (IoT)). Further drivers include a demand for reduced latency, real-time control, total coverage, improved quality of service (QoS), personalised service, and continued service while in high-density population areas or while travelling. Technologies include among others, extreme densification (small cells) and offloading to improve spectral efficiency; millimeter wave, multiple input multiple output (MIMO); orthogonal frequency division multiplexing (OFDM); software-defined networks (SDNs); network function virtualisation (NFV); cloud/fog computing; open transport protocol layer; cognitive RAT selection; device-to-device (D2D) communication; and several others. Most technology candidates for 5G era systems are aimed at the larger consumer market, however, these systems may extend to cover other use cases including emergency situations such as natural disasters and IoT-related applications such as driverless cars and e-health. Evans [81] argues the importance of the role of satellites in the 5G era, highlighting the improved coverage in rural areas, resilience and security they can provide in the event of a natural disaster, along with the locational benefit they can provide in the future of the IoT and intelligent transport systems. The current challenge lies in integrating the existing satellite technologies into the communications system architecture. The existing literature forms a consensus that the 5G era will see significantly more complex systems, and integration of those systems, than any previous generation of wireless technology, and will only be successful in meeting the set objectives as a result of multiple technologies combining to form a highly heterogeneous system.

The work by Osseiran et al. [202] identifies high-level objectives of 5G era systems: these include “Amazingly fast”, “Best experience follows you”, “Ubiquitous things communicating”, “Great service in a crowd” and “Super real-time and reliable connections”. For users with fixed broadband and WiFi access points at home, or for users served by operator-owned femtocells and picocells, a sizeable proportion of traffic generated by mobile and portable devices is offloaded from the mobile network onto the fixed WiFi network. As a percentage of total mobile data traffic from all connected devices, mobile offload is forecast to increase from 60 percent (10.7 exabytes/month) in 2016 to 63 percent (83.6 exabytes/month) by 2021 (Fig. 2.4) [21]. No clear boundary will separate WiFi and cellular from a users perspective in the 5G era. Services previously managed by one or the other will likely be available on both and the two systems will seamlessly integrate.

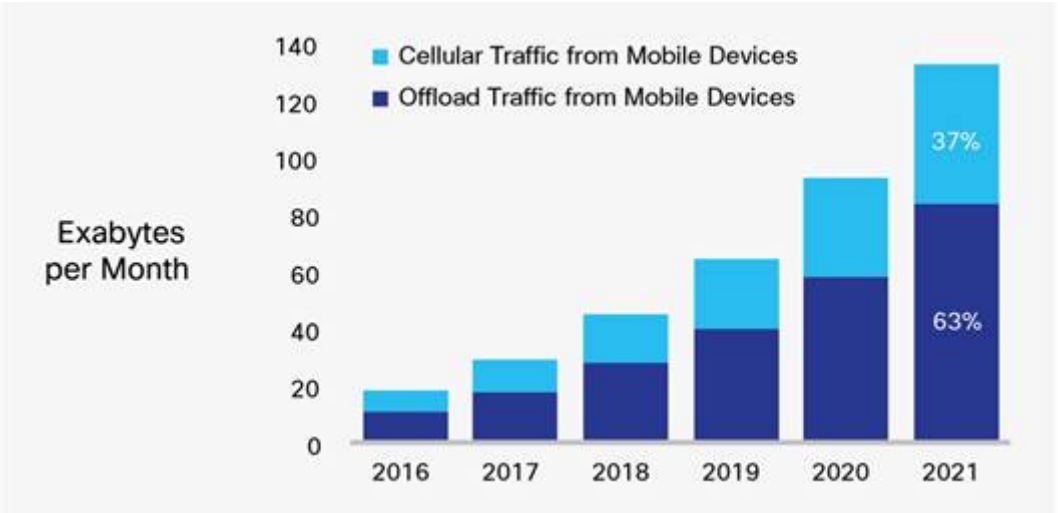


FIGURE 2.4. Global data traffic and WiFi offloading forecast. Figure reproduced from: T. J. Barnett, A. Sumits, S. Jain, and U. Andra, "Cisco Visual Networking Index (VNI) Update Global Mobile Data Traffic Forecast," Vni, pp. 2015–2020, 2015 [21].

2.3 Understanding New Business Opportunities for the 5G Era

Driverless cars, e-health, and virtual reality are all presently exciting areas of research with high-profile industry projects under way such as those by Google and Volvo [100, 261]. Other research grant-funded major projects include Venturer [257], UK Autodrive [255], FLOURISH [87], i-Motors [118], INTACT [120], Pathway to Autonomous Commercial Vehicles [94] and TALON [240], bringing together research institutes and automotive industry expertise, and further, UBIMON [254], CCHP [45], The Wearable Clinic [247], OPERA [198], CfPH [46] and SPHERE [234] working with clinicians, engineers, designers, and social care professionals. Proposed technologies in these areas rely on extremely high data rate transmission and other characteristics common of proposed 5G era networks.

Using driverless cars as an example: one mode of operation uses sensor and vision technology built into a normal car. The car could be released onto the road and observe traffic just as a human driver would. A further development to this model for driverless cars depicts a system whereby all traffic is constantly communicating with each other. Clever, centrally controlled routing algorithms automatically direct cars via different paths into a city centre to reduce congestion. Cars travelling fast on open roads are able to communicate and drive very close together, slipstreaming behind one another to make journeys more fuel efficient. At crossroads, cars approaching from all directions can communicate and calculate distances so precisely there will barely be a need to slow down; they will pass close to each other at speed. This kind of significantly more sophisticated autonomous transport requires high data rate transfer over a robust reliable network, which current communications infrastructure does not provide and

which therefore could benefit from 5G era capabilities. There are a number of projects under way investigating vehicle-to-vehicle and vehicle-to-infrastructure communication for infotainment and vehicle safety applications [75, 101].

Already e-health devices are available on the market to monitor daily activity levels for fitness and well-being applications. A reliable high-speed 5G era connection would allow for far more significant innovation in this area. For example, the SPHERE project [234] envisages that patients could be released from a hospital and remotely monitored in real time by doctors. Through monitoring vital signs, early warning of potential heart attacks or other emergencies could be achieved, improving the response of emergency services or even allowing doctors to give treatment to patients before symptoms arise. The aims of SPHERE are to solve real healthcare problems using technology which is acceptable in people's homes and to generate knowledge that will change clinical practice.

Virtual reality systems are emerging with major companies investing in the area. For example, in 2014 Facebook acquired Oculus VR for a significant sum [246]. It is possible to imagine a scenario of a virtual reality meetings, eliminating the need for business people to fly long distances to meet face-to-face. In order to achieve the desired life-like experience, a system would require extremely high data transfer, ultra-low latency and reliability. Current wireless networks do not have these capabilities but proposed 5G era systems may do.

The described examples depict existing ideas that will come closer to being achieved with the realisation of 5G era capabilities. The 5G environment will further provide a platform for new currently unforeseen innovations to flourish. These examples, and several others provide new business opportunities encompassed within the telecoms industry. There are opportunities for these new businesses to provide their own connectivity services or form relationships with existing telecoms industry operators or businesses to benefit mutually. The significant capital at the disposal of some of these companies enables them, if they wish, to threaten existing operators by investing in their own telecommunications infrastructure and system.

2.4 Problem Structuring Approach to Strategy Development for the 5G Era

2.4.1 A Complex Problem Context

This thesis considers the development of a strategy to address research priorities in the development of next-generation communications technologies as a *wicked problem* [214]. The following subsections explain how a systematic PSM has been used to manage the complexity of this problem and provide a structured approach to developing strategy that is likely to lead to business success in a 5G environment. However, first, the characteristics of a wicked problem and reasons why PSMs are a valid approach to take are articulated.

The definition of a wicked problem used is based in this chapter on the original 1973 work of Rittel and Webber [214]:

1. There is no definitive articulation of the problem situation.
2. Interventions have an impact on the problem situation, thus changing the context and leading to the need for further interventions.
3. Interventions in wicked problem situations can only form viewpoints that will regard them as improvements or a worsening of the problem situation, but not solutions.
4. The complexity of the wicked problem situation means that it is impossible to carry out experiments: interventions are one-off activities.
5. Many possibilities for intervention exist, including doing nothing; but articulating all possibilities may be impossible.
6. Problematic situations are likely to be part of wider wicked problems, and contain other wicked problems.
7. Data about the wicked problem, and evidence of change after an intervention may be contested.
8. Wicked problems require action to be taken to alleviate them: they are not studied purely for the purpose of knowledge gathering.

While highly cited and influential across many disciplines, more recent work from Mingers [178] summarises these principles as a more succinct set and more suited for use in Systems Engineering. The following embellishes his definitions with characteristics of the 5G era problem context to illustrate the relevant point:

1. Stakeholders in the problematic situation have different worldviews.

   Decision making about technology investments in a rapidly developing field such as 5G era communications is certain to be contested. Strategy development for an organisation requires achieving both shared understanding and shared commitment to action through a process of deliberation that is likely to start with disparate and possibly conflicting views.

2. There is no clear definition of the problem from the stakeholders.

   Sec. 2.2 illustrates something of the complexity of the 5G era technology development landscape. The multiagency nature of a 5G era service as a conceptual layer across multiple elements from multiple suppliers indicates a step-change in both technology and business models making the problem situation highly ambiguous.

3. The objectives of any intervention require agreement that is difficult to obtain.

   The existence of multiple routes to market for exploiting 5G era technology development poses a huge challenge to conglomerates to agree where best to direct effort to extract value from intellectual property (IP) development.

4. Definitions of success for interventions require agreement between stakeholders.

   Perhaps the most difficult aspect of strategic development is deciding among the stakeholders' measures of success. From the Rittel and Webber definition above, there is unlikely to be a clear finishing line for declaring a successful exploitation of any technology development in this space. Conventional measures of return on research investment are likely to be impossible to determine.

5. The problematic situation is characterised by high levels of uncertainty.

   While forecast and foresight methods will provide predictions concerning technology developments, the truth is that disruptive events may lurk in the future to disrupt the best-laid plans of any strategic decision-making process. Being aware of the fallibilities of projecting the past into the future must be part of any strategic decision-making process.

PSMs have been proposed as a valid response to dealing with such problems and there is wide literature on defining their characteristics, for example, Yearworth and White [279] and Ackermann [4] and an even broader evidence base of their successful use in many domains. Recent work has also shown how problem structuring can be seen as an essential activity in an enlarged scope for Systems Engineering practice [278]. This chapter follows the Generic Constitutive Definition (GCD) of PSMs from Yearworth and White [279] as an articulation of the key properties of a PSM and uses them in the design of the presented approach. This approach is described further in Sec. 2.4.4, but first the problem of modelling 5G era technologies and the integration of such models into the problem structuring approach is covered and also the question of who the stakeholders are is addressed.

2.4.2 Modelling 5G Era Technology Development

Mingers [177] summarises different techniques and modelling methods common to systems thinking; this chapter suggests where and how they could be applied in relation to the 5G era. This chapter provides a framework for categorising different techniques and indicates the philosophical viewpoint of each of the techniques explaining the ontology, epistemology, and axiology; that is, what they model, how they are modelled, and why they are modelled. This, in turn, indicates the stakeholders to whom this technique may be of interest.

This work implies if the technique is used to model physical (i.e., tangible items), social (i.e., the effects it will have on businesses and the economy), or personal (i.e., how a particular user will interact with it) aspects of the 5G environment. Consideration of the philosophical paradigm is

important in modelling the 5G environment, more so than with previous generations of wireless technology. Earlier generations of wireless technology were driven by performance-oriented goals of speed, latency, and coverage. This required a positivist approach, that is designing a system and showing, via testing and mathematical modelling, that goals had been achieved. The goals of the 5G era are much more stakeholder experience-driven (i.e., does it fulfil their needs to a satisfactory extent) and so require a more phenomenological approach, taking into consideration the needs of all stakeholders and understanding that these are unique. Different stakeholders, for example, mobile users, autonomous vehicles, e-health providers, and mobile operators will have a different measure to be satisfied by their experience of 5G era networks such as speed, reliability, the revenue it can generate, and others. By using a multimethodological approach [178, 179, 217, 218] combining the discussed techniques, all areas relating to 5G era systems can be covered. There are further techniques available not covered in Mingers work [177] and it may be possible to use some of the techniques in Mingers work for purposes other than that described which may change the underpinning philosophical assumptions. However, at a fundamental level, the mixing of such 'hard' systems modelling into a problem structuring approach is well understood and entirely feasible [146].

The models presented in this chapter were developed by group model building by a group of academic researchers (mostly post doctoral researchers with electrical, wireless and software engineering, and mathematics backgrounds) with expertise in a variety of technologies from all aspects of the telecoms industry and further the industries business infrastructure. The groups technical expertise included the PHY and MAC layers, and WiFi and cellular technologies. The models aim to use the expertise of the group in *developing a shared understanding of the 5G era development landscape* from which it is possible to draw inferences for the purpose of strategic planning for a business desiring to operate in the telecoms space as the industry evolves into a 5G era. The benefits of the modelling approach and what can be learnt from it are discussed below. From the models a list of 5G candidate technologies are identified.

2.4.3 Understanding the Stakeholders of the 5G Era

Fig. 2.5 demonstrates stakeholders' interest and their influence on the development of 5G era technologies. Furthermore, these are grouped by different coloured boxes roughly categorising them as regulator, business, demand, and technology drivers. All are contained within a larger box indicating the economy and environment.

The purpose of this stakeholder modelling is to make progress to understanding how their behaviours can be represented and analysed to help visualise abstract threats, opportunities, or other previously not considered issues in relation to the telecoms industry and how it will change with the launch of 5G era technologies and standards [264]. The capability to comprehend the often concealed power and influence of various stakeholders is a vital skill for success in a complex project. Stakeholders can be a significant asset contributing information, intuitions

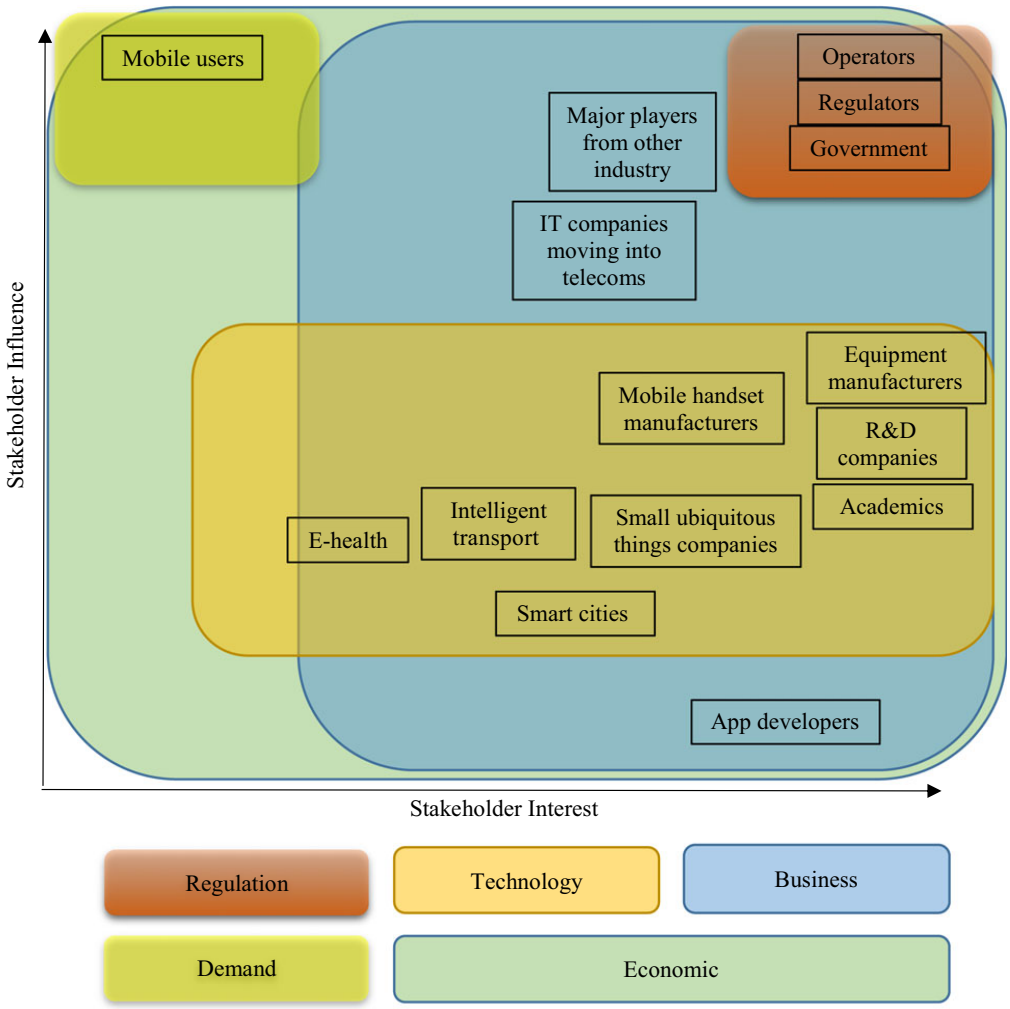


FIGURE 2.5. Fifth generation era stakeholder influence versus interest diagram.

and backing in defining a project and realising its execution. Any tools that help to identify and visualise stakeholders' are likely to impact and advance an organisations ability to address the complex problems and views of different stakeholders and their relationships [263].

All the regulating bodies fall in the top right corner of the diagram indicating they have high interest and high influence in the development of 5G era communication networks. Technology companies span the range of the interest scale but overall sit lower on the influence scale than regulators. No matter how innovative or brilliant a technology developed by an organisation, if the regulators chose not to include it in one of the many WiFi, cellular or other 5G era communications systems standards, it will have little impact. All the developments of technology companies are at the discretion of the regulating bodies. The regulating bodies are further embedded within the broader business category.

The market demand box falls in the top left corner of the diagram indicating low interest but

high influence. There is not presently great demand from mobile users asking for 5G technologies; however, the continual increase in demand for data due to the increasing number of connected devices and popularity of multimedia applications along with increased demand for improved coverage and reliability is a significant driver for the development of new systems.

Business encompasses a large area of the diagram; some businesses interested in the capabilities of the 5G era have very little influence on its development, other larger organisations hold significantly greater influence. A significant driver for 5G era capabilities is enabling new profitable businesses. Operators have seen a decline in revenue over recent years and see the 5G era as a new way of making money. Major players from other industries are looking to invest in telecoms as an opportunity to expand their source of income. New start-up businesses seek capital to develop their ideas on 5G era platforms. Together these stakeholders are changing the boundary of the traditional telecoms industry.

Mapping the stakeholder influence identifies the highest priority groups and which has the most power. In trying to determine what the 5G era will look like, it is perhaps most sensible to understand the vision of operators, regulators, and government who are identified in Fig. 2.5 as those with the highest influence. Mobile users are a high-influence but low-interest stakeholder. It is important during the development of 5G era technologies to engage with high-influence stakeholders. Approaching 2020 and the expected first release of 5G era cellular and WiFi networks, one should expect to see users move position in Fig. 2.5 further to the right. Those whose interest is already high such as operators must engage with users to ensure a satisfactory solution is reached. Consideration must be taken of those who fall in the bottom right of the diagram. Several key European projects that are developing 5G era technology and strategy have spent significant effort on identifying requirements through user-driven use case scenarios. Technology developers and academic researchers should be of particular interest to higher influence stakeholders since there is a chance their influence could significantly increase with new technology developments. The development of a disruptive technology could boost one of these to the top of the diagram. Those that fall in the bottom left of Fig. 2.5 should not just be disregarded. The best strategy for managing low interest, low influence stakeholders would be to provide them with information; this may increase their interest in moving to the right in the diagram, perhaps encouraging them to engage more with the development of 5G era technologies, bringing to light more information from which those with higher interest and influence can learn. Any business, existing or new, will need to engage with and understand the perspective of all stakeholders to best position themselves to be successful. An understanding of the role of other businesses / stakeholders, and how changes to their operations have an effect on the encompassing system, is essential to avoid systemic failure, a subject discussed in Boardman and Sauser [32] where they demonstrate how a complex system can collapse if one part fails. The authors go on to explain that to avoid this, each part must have an understanding and level of expertise concerning the role of each of the other parts such that should they fail, the gap in the

system can be filled and the system can continue to operate.

2.4.4 Problem Structuring Approach to the 5G Era

HPM was developed as a functional technique for modelling systems using a tri-valued representation of process performance based on interval numbers [66–68, 111, 166, 168]. The technique enabled very complex systems to be modelled using hierarchical process decomposition and the interval numbers allowed for explicit representation of process performance including epistemic uncertainty, that is, known unknowns. The value of the method has been demonstrated through application on a number of complex sociotechnical systems problems in domains such as an oil exploration project, flood defence systems, asset management, and performance management [67, 68, 111, 167].

Here, HPM was used to manage the complexity of problem structuring in order to understand how the transformational process <*improving our understanding of the 5G era development landscape*> could be achieved. This process was considered to be at the core of strategy development and the performance of the system to achieve this understanding would be affected by decisions concerning where to place resources. The process model, therefore, represents the top-level transformational goal as a composed set of subprocesses, structured into several layers to aid model development. Subprocesses are explicated through repeated application of a simple language game of asking how a process can be achieved and this continues until the successful performance of a process can be related to one or more of the technologies proposed for the 5G era. The model should be read as representing all the known processes necessary to achieve the transformational goal; in effect, the model can be viewed as a conceptual system that needs to exist to achieve the transformation. Higher level processes can be viewed as consisting of lower level processes in a part-of relationship. Although presented diagrammatically as a hierarchy, the 'sub' processes should be understood as being contained within their higher level process, not below. Each lower level is simply the higher level expanded. Fig. 2.6 demonstrates how the process works for the given transformational process. The technique is similar to the Purposeful Activity System modelling (PAS) that forms the core of Soft Systems Methodology (SSM) [50, 178].

Key drivers discussed are related via stakeholder drivers to the core technologies of the 5G era. The final diagram resulting from the group model building session can be used aiding understanding of what 5G era communications systems will encompass and further to identify analytically key technologies or aspects on which particular characteristics of the 5G era depend. The model evolved over several iterations and continued discussions with the group. The final model broke the highest level statement into three statements which represented the three main forces driving evolution to a 5G era: users, operators, and regulators. These were further broken down into lower-level processes until the high-level requirements could be linked to specific technologies. Additional annotations and subcategories are added to the model after the development process to help communicate its findings (see Fig. 2.7).

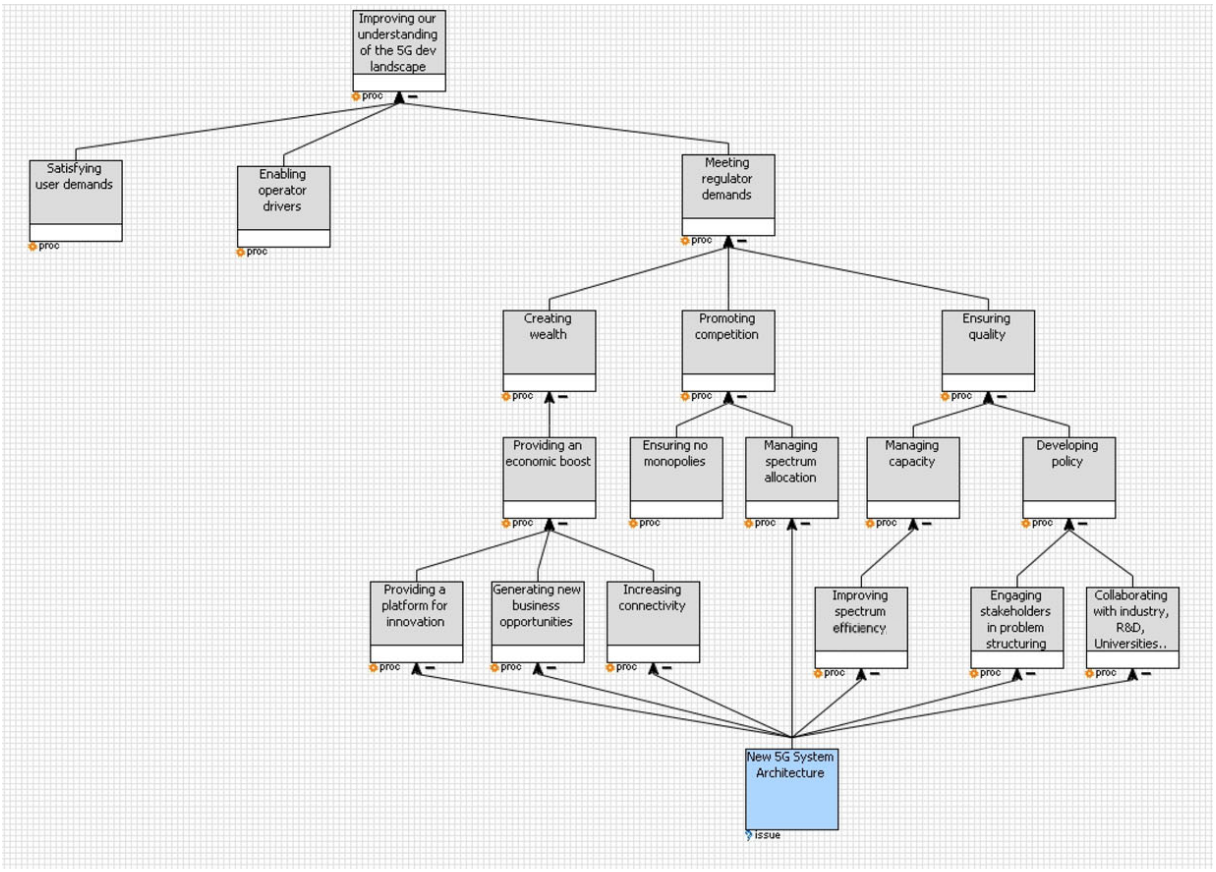


FIGURE 2.6. Hierarchical Process Model (HPM) under development representing part of the final model presented in Fig. 2.7. Higher level processes can be viewed as consisting of lower level processes in a part-of relationship.

The 'Technology requirements for 5G era' section of Fig. 2.7 could be considered as technical challenges and similarly the adjacent box as organisational challenges. The technologies appearing most frequently in the 'Technologies for 5G era' box could be thought of as the most essential. Where a technology only connects to one challenge, that technology may be critical to that challenge. It is likely that as time progresses, new technologies will be added and some may move position on the diagram. The diagram could be further expanded by further dividing the lowest level section, 'Technologies for 5G era', into their components. The diagram builds a conceptual vision of how all the proposed technologies of the 5G era are linked. 5G era systems will be complex socioeconomic-technical systems integrating different communications subsystems, using one or a few specific technology solutions and network architectures. The technologies and techniques identified in Fig. 2.7 could be implemented in numerous RATs including, but not limited to, cellular-based systems and further WiFi systems such as WLAN and capillary networks.

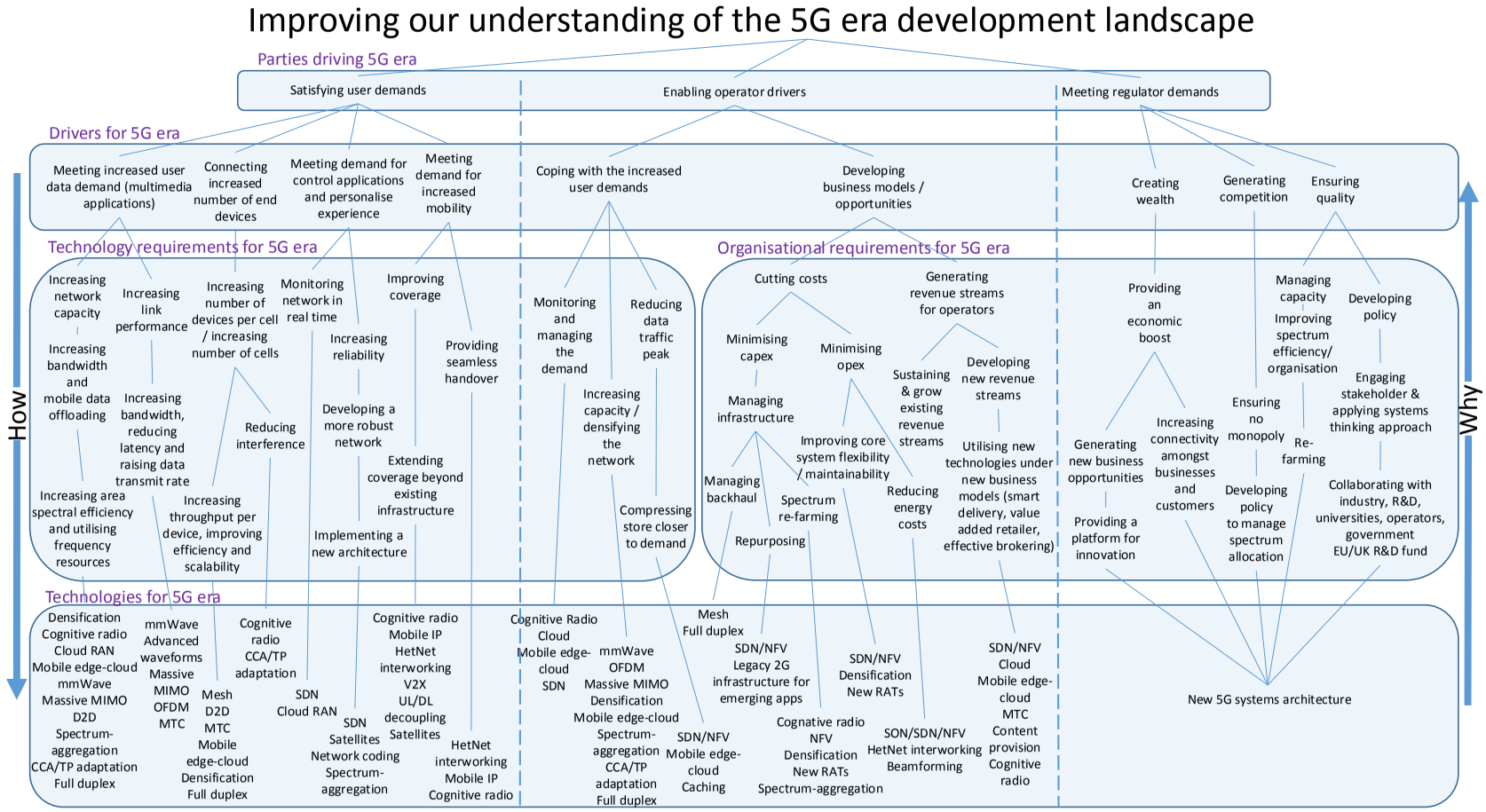


FIGURE 2.7. Hierarchical process model (HPM) final. Reading down the lower level processes explain how the higher level processes to which they are connected can be achieved. Reading up the higher level process explain why the lower level processes are necessary.

The application of HPM indicates the importance of different technologies to achieving the objectives of the 5G era. From the model a list of 5G era candidate technologies can be identified (see appendix. 2.A.2). The model shows some particular aspects of the overall network to be dependent on certain technologies; however, the overall indication it gives is of the need for all these technologies to work harmoniously to achieve the overall 5G era network. It clarifies the diversity of user, operator, and regulator demands and illustrates that the network will have to continually change and optimise for the changing demand and requirements. This indicates the need for an 'agile network' a phrase that nicely summarises the in-depth description of 5G era requirements from the perspective of multiple operators explained by the Next Generation Mobile Networks Alliance in their white paper [192].

2.5 Implications of Problem Structuring

The problem structuring is now discussed from the point of view of an equipment vendors' research and development lab whose interests centre around developing technologies to incorporate into future products or building an IP portfolio for potential licensing (i.e., Toshiba Research Europe Limited). The discussion takes a historical perspective and reflects on the future landscape of the 5G era.

Having developed a model via HPM, it is now possible to begin to use it analytically. From the model (Fig. 2.7), the first two columns of Table 2.2 (in App. 2.A.2 p.43) can be populated. The technologies are sorted in order of their HPM process connectivity score, that is, the number of connections a technology for the 5G era in the lowest layer forms with the layer above. All the technologies listed are proposed 5G era technologies and therefore may form a part of a 5G era system. Furthermore, now the third column of Table 2.2, '5G likelihood', is populated rating each individual technology as likely, very likely, or definitely. This uses the expert opinion of the same group that constructed the HPM to form a judgement on the current state of development of the technology and its adoption into the 5G era system, i.e., inclusion in one of the cellular or WiFi standards associated with the 5G era or inclusion in hardware / software products. The decision criteria is based on Technology Readiness Levels (TRLs) as defined by the Horizon 2020 Work Program 2016-2017 [115], however, it is simplified from nine categories into the three stated. For example, if a technology already exists and has been applied in existing communications systems, and is continuing to evolve, or is newly proposed with conclusive results demonstrating its capabilities, it will be marked 'Definite'. As a further example; if a technology newly proposed for the 5G era is in development with indications that it will provide significant advantages for 5G era systems but has yet to be demonstrated entirely conclusively or has come up against some opposition in terms of inclusion into standards, it may be marked 'likely' or 'very likely', depending on the expert opinion of those compiling the table and their judgement as to whether the technology will be ready for inclusion in 5G era standards. The judgement thus forms a

relative scale between the technologies. This process further helps to address any conflicting requirements from the HPM by considering the influence of the stakeholders in seeing their requirement realised.

This chapter has used a PSM to facilitate exploring the question of how to develop a strategy for research investment in the complex development of 5G era technologies. In the earlier review of the historical development of the telecoms industry, situations were identified where lack of strategy led to the decline of companies. Through mapping stakeholders and categorising them (Fig. 2.5), the significance of a business push underpinning to stakeholders is recognised. By application of HPM, the business drivers of key stakeholders are related to technologies. To continue building an understanding of the 5G era landscape, it is necessary to understand how businesses manage the technologies to provide business gain in order to identify which technologies will prove popular among the stakeholders of the 5G era. Therefore a second iteration of the problem structuring methodology is performed to explore the original question encompassing the areas highlighted as requiring further research, as outlined in Sec. 2.6. An understanding of the 5G era landscape is developed by exploring the existing and proposed business models in the telecoms industry.

2.6 Developing Understanding of the 5G Era Business Landscape

Systems engineering principles ensure stakeholders engagement is a priority throughout the development process of any new design, technology, or system. A systems engineering approach to conducting business in the 5G environment assists in recognising the social, policy, and business challenges involved, as well as the technical engineering challenges. At present, there is very little in the literature regarding business models for the 5G era other than speculation. Any stakeholder wishing to be successful in developing or operating in a 5G environment needs a business model, created in parallel to technology progression, within the bounds of regulatory requirements and their progression. Therefore, the overriding challenge for any stakeholder is to create a sustainable business infrastructure encompassing technologies of the 5G era, and accounting for potential future innovations, such that they can be profitable while providing a high-quality product or service at an affordable cost to their clients [202].

2.6.1 The Historical Context: Understanding Challenges to Traditional Business Models

The traditional business model of a mobile network recognises operators owning infrastructure and spectrum, and then charging subscribers for using it; this model has remained largely the same through the first four generations of mobile technology. Today mobile network operators face major challenges [195]. They are under pressure with revenues declining worldwide. Reasons

for the decline include: first, a saturated market forcing down prices and therefore revenues; second, significant expense to operators is created keeping up with growth in demand due to the explosion in data traffic; and third, operators have been slow to adapt their structure for operational expenditure relative to a rapidly evolving market. Furthermore, the impact of the worldwide economic downturn of the last decade has added further pressure on operators.

Traditional business models are also put under further pressure by changes to the wider industry such as regulation amendments, changes in enterprise behaviour, new technology, and changes in consumer behaviour such as increased use of free communication apps taking revenue away from the operators while still using their network capacity. This kind of change in user behaviour is a social change and a driver for change in operator business models.

As discussed in Osseiran [202], companies previously considered to be IT or data-communication organisations are venturing into what was previously considered to be the telecom operators domain, having seen a market for functions that smartphone users could benefit from, such as cloud-based services. This puts further pressure on telecoms companies to provide a similar service at a comparable cost in addition to the standard expected cellular and data service. Upgrading the network is likely to be the biggest cost to the operator.

Furthermore, business and technology collaborations or developments might see companies outside of the telecoms world challenge or partner existing operators [202] by investing in their own cellular spectrum and data services. With the expansion of the 'Internet of things' and emerging technologies such as the intelligent transport business, giants from other industries may wish to invest either to own spectrum or in infrastructure to provide services to mobile users without depending on traditional operator networks. New niche business models face great challenges in meeting the speed, reliability, and robustness goals of the 5G era that users will expect at an affordable cost. Current telecoms operators may counter these and adapt their own business models, contributing their experience and expertise to developing new services and solutions in application domains such as transport and energy, through partnerships with large service sector players in transport and utilities.

2.6.2 Understanding State-of-the-Art Business Models

While clearly major telecoms operators will continue to play a major part in management and operation of the network, potential new major customers or potential providers such as the automotive industry and those developing 'Internet of Things' type devices, and how they will shape the 5G era landscape, should be considered.

Looking at previous mobile technology generations, there has not been much evolution in the business models since cellular technology turned digital [40]. For users and operators, each new generation is simply a faster version of the previous. The fundamental differences in the architecture of 5G era systems [164] will require considerable changes to standardisation and to the business model.

The functions of a 5G era business model are:

1. Express value, that is, what is the value created for a specific business by its 5G era capabilities and the new advanced features it offers.

2. Identify a market segment, that is, who are the users of the 5G era related service and what is provided?

3. Define the structure of the value chain required by the operators, regulators, and users to create and distribute the service, and determine the complementary assets needed to support their respective position in this chain. Each stakeholder must consider their suppliers and customers, and their view of the system should extend from resources to the final customer.

4. Specify the revenue generation mechanism(s), and estimate the cost structure and profit potential given the value proposition and value chain structure chosen.

5. Describe their position within the 5G environment, holistically linking suppliers and customers, including identification of potential complementers and competitors.

6. Formulate the competitive strategy by which the innovations of the business will gain and hold an advantage in the 5G environment over competitors and previous strategies stemming from earlier generations of communications technology [55].

Some of the proposed technologies of the 5G era found in Felita and Suryanegara [84], Gohil et al. [97], Andrews et al. [10], Hossain et al. [116], and Osseiran et al. [202] already exist; the challenge in realising 5G era requirements is about developing an architecture for linking them together with the new technologies.

Many goals outlined for the 5G era could be achieved by simple improvement of 4G era systems. For example, complete nationwide coverage could be achieved by improving the current infrastructure and high speed could be achieved by improved spectrum optimisation. Technology can be viewed as a way to fulfil a purpose as explained by Arthur [12], who explains that new technologies are not always inventions that come out of nothing; rather most new technologies are constructed by combining existing ones. This definition would suit the development of wireless communication, since existing technologies being brought together to realise a set of goals will see the realisation of a new technology. Evolution of the 4G era and the incorporation of new technologies and stakeholders will form the 5G era.

In the modern world, businesses are recognising the power of wireless connectivity as a platform for innovation. This thesis chapter argues that the 5G era will impact the process of innovation, a key procedure in the creation of value. 5G era technologies will provide a platform for interactivity. Business engagement with customers, robust reliable communication, and

increased speed and flexibility will generate novel ways of generating revenue [10]. Companies can use the capabilities of the 5G era to engage customers in collaborative innovation.

Nokia Siemens Networks identifies three strategies mobile operators or other businesses might pursue for a 5G era depending on their capabilities and market conditions [195]:

1. 'Smart delivery': providing user-specific cleverly managed services to generate additional revenue opportunities through collaboration with content providers and global service operators.

2. 'Value-added retailer': by considering each user as unique, and by effectively designing their service to reflect this, value can be added by improving the QoS.

3. 'Effective brokering': by monopolising on the capabilities, knowledge and assets existing operators have available, such as customer loyalty and market awareness. It may be possible to occupy a brokerage role, bridging a gap between mobile users and businesses that may mutually benefit each other.

Some aspects of proposed 5G era systems will likely be incorporated by a simple adoption of an operator's existing business models [202]. However, 5G era systems will see significant technical advancement from previous generations. This state-of-the-art-type technology needs to be matched by state-of-the-art business techniques. The strategies posed by Nokia Siemens Networks illustrate the sophisticated business acumen required to support the sophisticated technology. The proposed strategies further extend beyond operators to others looking to profit in a future 5G environment. 'Smart delivery' is something app developers must consider when designing their service to make it user-specific. Many IoT companies very much fit the description of 'value added retailer'; by collecting data they become unique to their user and provide feedback or benefits. As new big players, such as those from the automotive industry who look to gain from the 5G era, effective brokering, in terms of both assets and customers, could prove beneficial. The second largest U.S. mobile network operator by number of users, AT&T, previously announced that they consider connected cars as one of the next big growth opportunities [245]. The automotive industry and telecoms operators will both be looking to develop business models and investment strategies that best position each of them and their respective customers as this market develops.

The three proposed strategies are closely interlinked and are strengthened by each other. For example, a value-added retailer strategy may be applied to develop the best QoS for the customer and then smart delivery techniques may be applied to deliver that service in the most cost-effective way to the operator thereby maximising the operator profit. They may work together perhaps either via one large company in different divisions or by more specialised companies collaborating.

Methods of charging and billing customers will likely diverge from existing methods. With the increasing number of free communication apps and free WiFi access points, along with emerging

technologies such as D2D providing communication without going via the spectrum, operators will need to provide good incentives, that is, very high QoS, for users to pay for their services. Other businesses may look for alternative revenue streams from users with connectivity as a free service to attract customers.

A principal analyst with Quocirca, the business and research house [98] speculates that usage costs will move on from call time and data speed calculations to layered speed-based charging. He states that a universal basic (relatively low throughput) service may be offered, at no charge or minimal cost for all devices, with the option of additional tariffs as the data speed goes up. 5G era services are likely to be available to everyone on a speed-based pricing model that allows users to choose which speeds they wish to pay for.

Yazici [277] further discusses the idea of a layered model proposing the Network Controller on top of the pyramid. Various operators could apply this business model as part of a software-defined architecture that would govern QoS provisioning from operator to user, application-aware routing, user mobility, political issues, and revenue generation, delegating to several lower level control processes. The architecture proposed by the authors introduces an integrated service methodology to mobility, handoff, routing management, and connectivity management. Organised by complex control plane functions, this architecture also suggests a speed or data flow based charging system.

2.6.3 Applying Problem Structuring-Systems Analysis of Business Models to Further Aid Understanding

Tab. 2.1 (p.40) summarises potential business models that may emerge during the evolution of the 5G era. The information summarises content presented earlier in this chapter and is further influenced by Nokia [195], NGMN Alliance [192], China-Mobile [57], and KPMG [147].

Methodologies discussed in Boardman and Sauser [32] can be used to map the potential new business models and their routes from drivers for change to the telecoms industry. Fig. 2.8 shows diagrammatically how demands on next-generation networks drive either the continued development of the traditional operator-owned infra-structure service model or new business models (smart deliver, effective brokering, and value-added retailer).

The process of reviewing the telecoms business literature and mapping the drivers for the development of the business models further aids in developing an understanding of the landscape of the 5G era and achieving more informed decision making in relation to a strategy for research investment. The process provided an understanding of how technology areas in which one may wish to invest research effort might be utilised. Despite the potential benefits, a technology that cannot be incorporated into a business model is unlikely to be widely used and therefore may not be a sensible investment of research resources.

With improved understanding of the 5G era business landscape, it is now possible to further develop Tab. 2.2 (p.43). Improved understanding enables the fourth column of the table, ‘Action’,

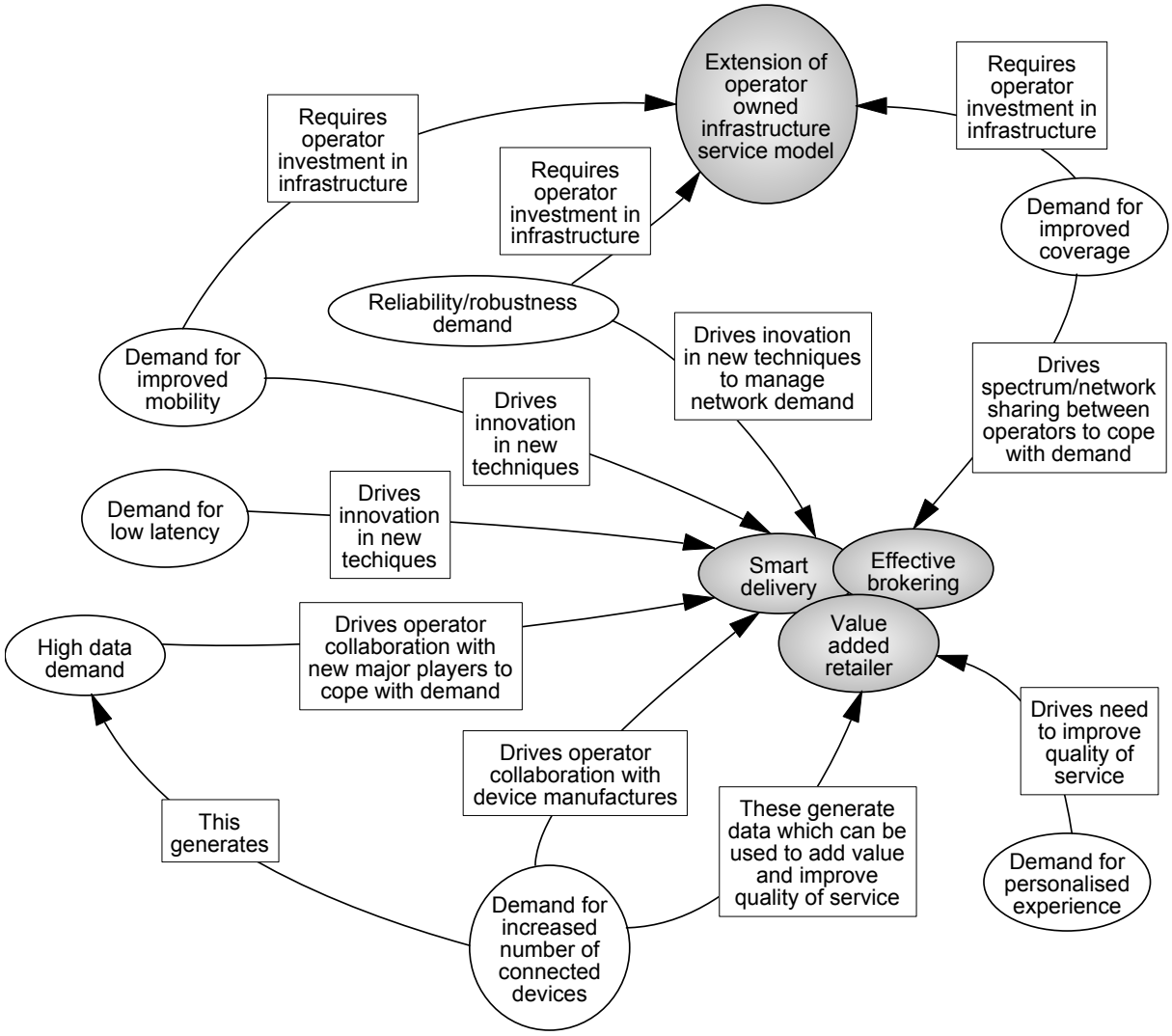


FIGURE 2.8. A systemigram illustrating the link between the drivers for the 5G era and the development of business models.

to be completed. Each technology is assigned a letter A, B, C, D, or E according to the key at the bottom of the table. Furthermore, the comments column of the table is also populated where further explanation or justification for the decision is needed.

2.7 Reflections on 5G Era Strategy Development

Establishing a process such as that recommended in this chapter minimises the risk of strategic failure due to lack of understanding of the surrounding system [32]. Through continued iterations, the process will remain largely the same, however, as the surrounding system develops the models will change to reflect that. In this situation, knowledge and understanding of the benefits of an established process are as important as the content knowledge.

The iterative process presented in Fig. 2.9 includes the loop reviewing the history and state-of-the-art. It may seem at first that these will only need to be done at the beginning of the process; however, this is not ideal for the following reasons. The initial historical review, as presented earlier in this chapter, provides context and understanding. In an industry of the scale of the telecoms industry, a review of the entire history in detail is not feasible. Instead, the best approach is to use some expert judgement and investigate areas thought to be significant. Moving further through the process presented in Fig. 2.9 will highlight other areas as important. Further exploration of the history in relation to these newly identified important areas may be required in additional iterations of the loop. Similarly, after reviewing the state-of-the-art, the problem structuring may further highlight areas that were not included in the first review. Furthermore, the time scale of the loop is not specified. Continued iterations may take place over sufficient time such that the industry state-of-the-art has moved on: this chapter presented the first iteration. In addition to this iterative process, to make an initial strategic decision, the whole process will need to be repeated periodically to confirm that the strategy selected is still suitable and recognise any amendments that need to be made. There is no set time interval at which this should be done; a decision will have to be made taking into account the rate of progression in the industry, and in response to any changes in the industry landscape. The landscape developed from the first iteration will help users of this methodology make the decision concerning how often it should be revised.

In relation to 5G era strategy, the process highlighted areas where there is an opportunity for a research lab to make technical contributions with a supporting business case. The process provided an understanding of the technology and business push for 5G era capabilities and an understanding of how these two drivers interrelate. The resulting understanding from the application of this process enables more informed decision making, thus reducing investment risk.

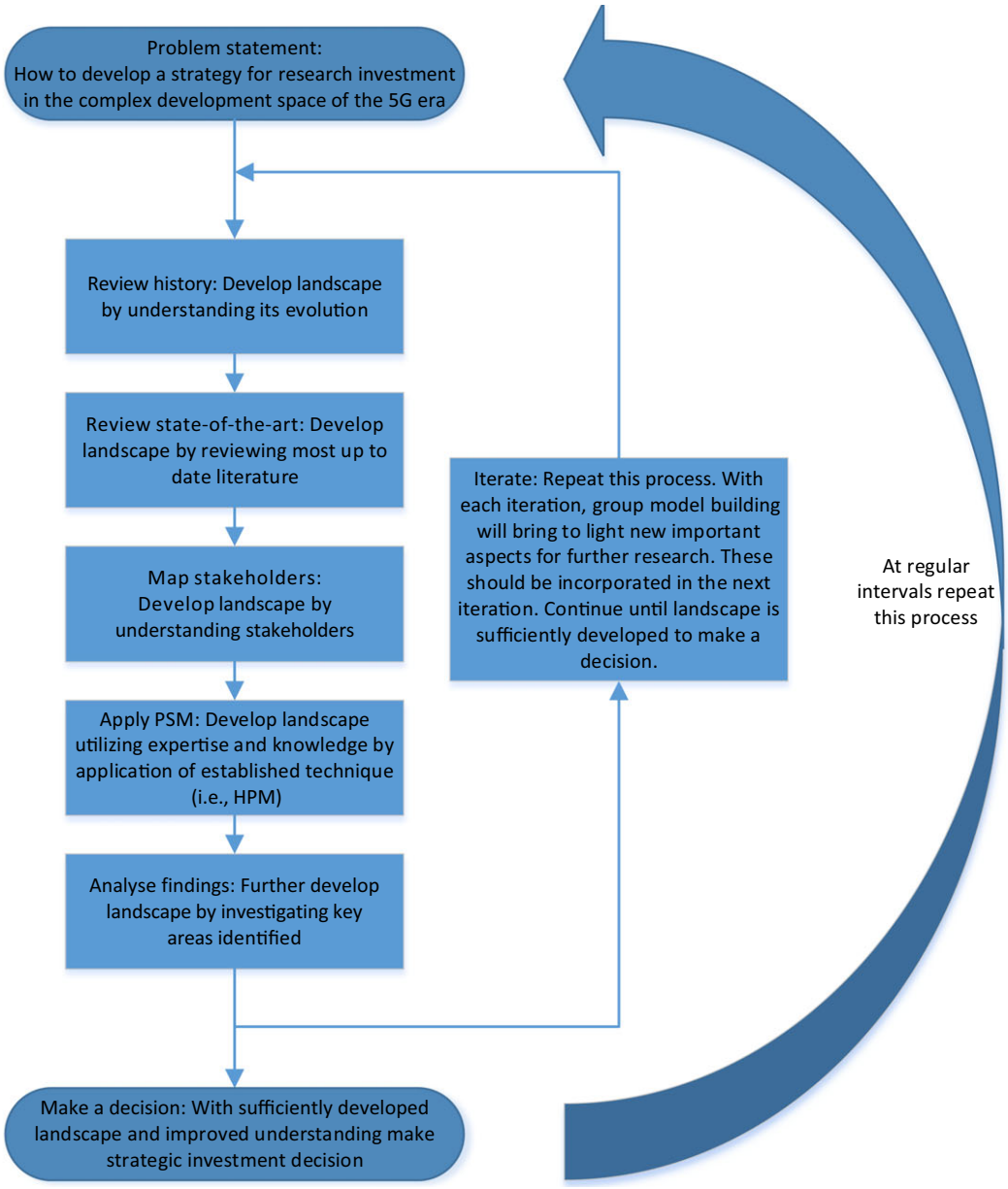


FIGURE 2.9. The methodology of this chapter; from problem statement to decision making. This chapter presents one iteration of this loop. Key Terms: 5G - Fifth-generation communication systems; PSM - Problem structuring methods; HPM - Hierarchical process modelling.

2.8 Conclusions and Scope for the Remainder of the Thesis

This chapter contributes to knowledge in the following ways:

CTK 2.1 Presenting a methodology for developing a strategy for research investment in the complex development space of 5G era technologies.

CTK 2.2 Demonstrating how specific technical drivers of the 5G era are causing the evolution of the business infrastructure.

CTK 2.3 Demonstrating how problem structuring techniques can be used to develop a vision of the landscape of the 5G era for more informed decision making.

The chapter explored the application of systematic PSMs to the development of research strategy in response to the challenges of the 5G era. It is hoped that the background provided will have helped develop an understanding concerning how the telecoms industry has evolved historically to reach its present state and how the business side of the telecoms industry is evolving toward the 5G era, the drivers for it, its complexity, and the state-of-the-art next-generation technologies that are involved. The remainder of this thesis aims to provide further insight into some of those identified next-generation technologies, providing quantitative analysis of the potential gains, and demonstrating effective strategies for implementation. The wider ongoing development of the industry will continue to be monitored and the 5G era landscape that has been developed will evolve accordingly.

As cellular and WiFi have evolved, the two have become increasingly interwoven in modern telecoms systems, driven by the need to accommodate the significant and continually increasing number of connected devices (see Fig. 2.3 p.16), and the ever-increasing demand for data (see Fig. 2.4 p.19). Offloading data traffic to WiFi has been a key strategy employed by operators and the proportion of traffic offload is expected to continue to increase to 65% by 2021. This thesis, therefore, focuses its investigation of the identified key next-generation technologies (see Tab. 2.2 p.46) within the context of WiFi networks, specifically IEEE 802.11 CSMA/CA networks, the most widely adopted WLAN technologies. It is intended that the approach presented for investigating next-generation technologies' impact on CSMA/CA networks is transferable to other RATs.

The drivers discussed in Sec. 2.2.3 (p.15) present more than just a technical challenge. The problems, or goals, are not clearly defined. The requirements are broad, open-ended and dependent on the complex changeable environment in which the networks are contained; subject to huge variation in user density and traffic, with considerable difference between the peak and the average data demand. All the requirements must be satisfied, yet, using existing technologies, achieving some may impact negatively on others. Developing solutions requires using research to build appropriate understanding with methodologies capable of modelling these complex problems for which there is a shortage of data available.

These requirements outline a significant research and development challenge to the wireless industry. The changeable environment of the wireless medium, the high user density and seemingly conflicting requirements, such as increased throughput and reduced energy consumption, need to be quantified with methodologies capable of capturing the relationship between these factors. In particular, models capturing the detail and complexity of next-generation features at the level of individual links, simulated such that the impact on the wider network can be observed, are necessary for improved understanding and design.

2.A Appendix

2.A.1 Summary of Business Models

Tab. 2.1 summarises the business models of the 5G era introduced in Sec. 2.6. The table relates key stakeholders to potential future business models and discusses the opportunities and potential threats the new models pose.

Table 2.1: Summary of business models

Basic Description	Key Stakeholder	Detailed Description	Business Model	Positives / Negatives of Proposed Business Model
Evolution of 4G era	Operators, regulators	Operators continue to incorporate new technologies, maintain and upgrade the current systems. Higher charges may apply for faster services.	Same as present 4G era; service model	As demand on mobile networks increase operators will not be able to afford the required improvements to infrastructure. There is nothing to stop others moving into the telecoms industry and undercutting current operators.
Network sharing / collaboration / over the top	Operators, service providers	Operators and service providers share physical infrastructure and spectrum and charge each other for usage. Controlled by SDN	Advanced service model incorporating smart delivery and effective brokering	For some operators or service providers, this will make it difficult to distinguish themselves and may take away key features they use to attract customers. Alternatively this can reduces the risk operator's face by reducing the need to continually invest in depreciating infrastructure reducing capital costs. Further it will improve efficiency and may help manage the difference between peak and average demand.

Table 2.1: Summary of business models

Basic Description	Key Stakeholder	Detailed Description	Business Model	Positives / Negatives of Proposed Business Model
IT cloud servicing companies expand to telecoms	IT cloud service providers, operators, regulators	IT companies may collaborate with telecoms industry (partner service provider)	Smart delivery	Opportunity for mobile operators to use strengths of IT companies to achieve improved QoS. Also may help to reduce or delay need to upgrade infrastructure.
		IT companies invest in telecoms spectrum (connectivity provider)	Wireless communication becomes a utility	Major threat to undercut pricing of mobile operators.
Other industries enter telecoms industry	Major players from other industry, for example, automotive, aeronautical, satellite, etc., operators, regulators	Major players in other industries collaborate with mobile operators (partner service provider)	Smart delivery / effective brokering	A chance for operators to collaborate with new market or other industry sectors such as transport which has increased demand for communications not only for passenger connectivity but to implement services based on various types of data (monitoring and control data for tracking, safety, security, control, management, etc.).
		Major players invest in their own spectrum and telecoms infrastructure (asset provider)	Wireless communication becomes a utility	Major threat to undercut pricing of mobile operators.

Table 2.1: Summary of business models

Basic Description	Key Stakeholder	Detailed Description	Business Model	Positives / Negatives of Proposed Business Model
Small companies entering telecoms	Smart technology companies, E-health, ubiquitous things, etc., operators, regulators	Ubiquitous things communicating, novel inventions by small companies may wish to use wireless resources	Value-added retailer / effective brokering	An opportunity for operators to partner new companies, perhaps some sort of system of leasing resources, to allow new companies to rise up.
New major player moves in	New major players, operators, service providers, regulators	A new major player moves in offering very cheap (or even free) service in return for users. A possibility if changes are made to the regulation surrounding the unlicensed spectrum opening it to use by anyone. Technologies such as licence assisted access could be utilized.	Very low cost or free to use. Provider benefits from increased user numbers (requires major initial capital, e.g., Google)	Completely undercuts traditional operator charging. Likely to result in huge reduction in customers for traditional operators. Very disruptive technology.

Table 2.1: Summary of business models

2.A.2 Decision on Research Investment

Tab. 2.2 summarises the decision on research investment. Further to problem structuring (Sec. 2.4), the table can be populated with 'HPM score' , that is, the number of connections a technology for the 5G era in the lowest layer forms with the layer above in Fig. 2.7. Furthermore, the third column of the table '5G likelihood', is populated rating each individual technology as likely, very likely, or definitely. This uses the expert opinion of the same group that constructed the HPM to form a judgement on the current state of development of the technology and its adoption into the 5G era system, i.e., inclusion in one of the cellular or WiFi standards associated with the 5G era or inclusion in hardware / software products.

With additional exploration of the 5G era business landscape (Sec. 2.6), it is further possible to populate the fourth column of the table, 'Action', to be completed. Each technology is assigned a letter A, B, C, D, or E according to the key at the bottom of the table. The 'comments' column of the table is also populated with explanation or justification for the decision.

Since undertaking the group model building process that created the HPM and this list of technologies, developments in the wireless industry have occured which may impact the list. More standards relating to the 5G era have emerged and technologies have further developed such that it is now clearer whether or not they will form part of any 5G era system. For example, developments in mmWave and Full-duplex would likely change the related status in the '5G likelihood' column of this table from 'Likely' to 'Definitely'. Similarly, if the modelling process were to be repeated, some technologies may have appeared in more or fewer places than in this iteration. This, hence, would have impacted their ranking in the following table. The HPM (Fig. 2.7) and this table represent one iteration of the methodology summarised in Fig. 2.9.

Table 2.2 Decision on research investment strategy

Technology	HPM Score	5G Likelihood	Action	Comments
Mobile edge cloud	10	Definite	A	A broad area with significant opportunities.
Software defined networking	8	Very likely	A	Key to smart delivery.
Network function virtualisation	6	Very likely	A	Key to smart delivery (extremely closely related to SDN, there HPM score could be considered combined).
Cognitive radio	7	Likely	B	The benefits in this seem very significant and it could play a significant part in smart brokerage-type business models.
New 5G system architecture	5	Definite	C	This still has significant unanswered technical questions and no standard.
Densification	5	Definite	A	Fits all proposed business models and important for handling increasing demand. Historical trend utilising small cells only seems to be accelerating. Will require sophisticated interference management techniques to enable high density networks to operate.
Spectrum aggregation	4	Definite	A	Thought to be a significant disruptive change to the industry. Many opportunities for innovation.
Full duplex	4	Likely	B	Significant challenges remain in realising full duplex however the benefits are high and there are many applications.
mmWave	3	Very likely	B	By many PHY specialists consider the key ingredient for the 5G era, a significant community pushing its development and inclusion in standards, yet to be conclusively tested.

Table 2.2 Decision on research investment strategy

Technology	HPM Score	5G Likelihood	Action	Comments
Massive multiple input multiple output (MIMO)	3	Very likely	A	A continued evolution of a well- established technology. Fits historical trend. Only question is where the line will be drawn in its development by standardization. Its evolution will continue beyond the 5G era.
Heterogeneous network (HetNet) interworking	3	Definite	A	Heterogeneity has been a clear trend in the wireless industry. Closely related with densification and could be considered the same research area.
Machine type communication (MTC) otherwise known as machine to machine (M2M)	3	Very likely	B	Many opportunities to profit from this technology outside of core communications systems. Related to others in this list, for example, V2X.
CCA/TP adaptation	3	Likely	A	Clear opportunity's for development. WiFi / cellular standards are not likely to give much constraint on this so plenty of opportunity for innovation.
Cloud radio access networks (RAN)	2	Likely	C	Significant unanswered technical questions for it to be realised in the 5G era.
Device-to-device (D2D)	2	Likely	C	Opposed by some operators and regulators as reduces their control of the network.
Orthogonal frequency division multiplexing (OFDM)	2	Definite	D	Quite a mature technology.
Mesh	2	Very likely	B	Risk is minimised as skills and knowledge are transferable to other applications.
New radio access technologies (RATs)	2	Definite	B	Many proposed options here. Further analysis will be required to identify the role they will play in the 5G era.
Cloud	2	Definite	D	There is significant competition from other out of industry bodies (i.e., IT giants).

Table 2.2 Decision on research investment strategy

Technology	HPM Score	5G Likelihood	Action	Comments
Satellites	2	Likely	D	Seems quite out of our area currently but as technology develops and integrates in to core system opportunities may come available.
Mobile internet protocol (MIP)	2	Definite	D	Mature technology. Number of variants exist, for example, HMIP, PMIP.
Vehicle to everything (V2X)	1	Very likely	B	How this will integrate with 5G era systems is unclear but technology developed here can be utilised in other ways, i.e., with the development of the automotive industry, particularly autonomous vehicles.
Upload (UL) Download (DL) Decoupling	1	Likely	C	Technology in early stages. May provide opportunities in the future.
Advanced waveforms	1	Definite	C	Likely that opportunities may emerge.
Beamforming	1	Definite	D	Mature technology.
Network coding	1	Likely	C	An area that needs better understanding before it will be included into 5G era standards.
Caching	1	Very likely	B	Opportunities for IP and uses in other areas: WiFi, M2M, sensor networks.
Legacy 2G infrastructure	1	Definite	D	Seems quite out of our control.

Key:

A. Something to invest time and research effort into. There is opportunity to develop products or intellectual property in this area. It will play a role in the 5G era.

B. Something to invest time and research effort into. There is opportunity to develop products or intellectual property in this area. It is likely to play a role in the 5G era.

C. Something to monitor. There is opportunity to develop products or intellectual property in this area. The technology may not definitely play a role in the 5G era.

D. Something we should monitor. Currently, cannot see an opportunity to develop products or intellectual property in this area.

E. Something to ignore.

Table 2.2: Decisions on research investment strategy

CHAPTER 3

A Comparison of Modelling and Simulation Research Methodologies for Wireless Networks and Discussion of Their Suitability for the Development of the Next Generation of Wireless Networks

This chapter defines and proposes a conceptual simulation methodology and argues why it is a useful approach for research in the development of CSMA/CA wireless networks. Some background on CSMA/CA networks is provided and further extensive literature review of modelling and simulation research in relation to them is presented. Following examination of the existing literature, categories of research methodology, namely analytical modelling and normative simulation modelling, are compared with the proposed conceptual simulation modelling approach. The inherent assumptions of these methodologies are discussed, in parallel with their suitability and potential contribution to the development of next-generation wireless networks. The methodologies' capabilities are discussed. It is identified that a validated conceptual simulation modelling methodology uses the strengths of problem structuring methods to overcome limitations of purely analytical techniques.

This chapter is structured as follows: firstly Sec. 3.1 provides background on CSMA/CA networks. Then Sec. 3.2 reviews existing modelling and simulation approaches in relation to wireless networks. Further Sec. 3.3 identifies opportunities for innovation in relation to modelling and simulation, proposing a new conceptual simulation modelling approach and comparing it with existing methods. Finally, Sec. 3.4 gives conclusions.

3.1 Background: Wireless CSMA/CA Networks

Wireless communications signals transmitted simultaneously may mutually interfere with each other (assuming that all transmitters are using the same channel). The Institute of Electrical and Electronics Engineers (IEEE) 802.11 Distribute Co-ordinate Function (DCF) CSMA/CA MAC protocol [119] is designed to coordinate transmissions between stations, such that they can share access to the physical medium fairly. Basic access CSMA employs a simple listen-then-wait system where transmitters wait for the physical medium to be silent before speaking. With this mechanism, there is no assurance that the transmitted message has been received at its intended destination. The CSMA/CA mechanism extends the listen-then-wait protocol to include a confirmation that the transmission has been successfully received.

Multiple channels can operate in parallel across the physical medium, however, the demand level is such that it is important to maximise the performance of each channel. Thus for simplicity, the following protocol description and the research throughout this thesis assumes all transmitters and receivers share one physical channel.

The following background sections provide the necessary technical understanding of the method by which the CSMA/CA MAC protocol achieves successful transmissions between nodes and minimises collisions. It is directed to non-experts in the MAC protocol so that they may understand the contributions later in the thesis. In Sec. 3.1.1 the MAC protocol is described in layman's terms along with the mechanisms by which it may fail. Fig. 3.1 illustrates a successful CSMA/CA transmission and further Fig. 3.2 illustrates the relationship between each of the described stages of the protocol, the conditions that must be met for the stage to be successful, and the mechanisms by which it may fail. Sec. 3.1.2 further explains the key backoff mechanism employed by the CSMA/CA protocol to minimise collision and manage contention between competing demands (see Fig. 3.3). Sec. 3.1.3 then explains how when implementing the MAC protocol, the physical environment may affect the performance (see Eqs. (3.4) and (3.6) p.56 and associated explanation) and further introduces the classic wireless network problems of hidden and exposed nodes.

3.1.1 Layman's Guide to the CSMA/CA MAC Protocol

In order to understand the CSMA/CA protocol, it is best to imagine a conversation between two persons named A and B. Person A wants to send a long message to B. It is vital to A that he knows that B receives the message in its entirety and without misunderstandings / errors. In particular, if a third person, C (say), were to start talking at the same time, then B might not hear A correctly, and the communication would fail.

To solve this problem, the CSMA/CA protocol employs handshakes between A and B. A successful communication, expressed in layman's terms, proceeds as follows:

S1. A listens to determine if anyone is already speaking. This is called *carrier sensing*.

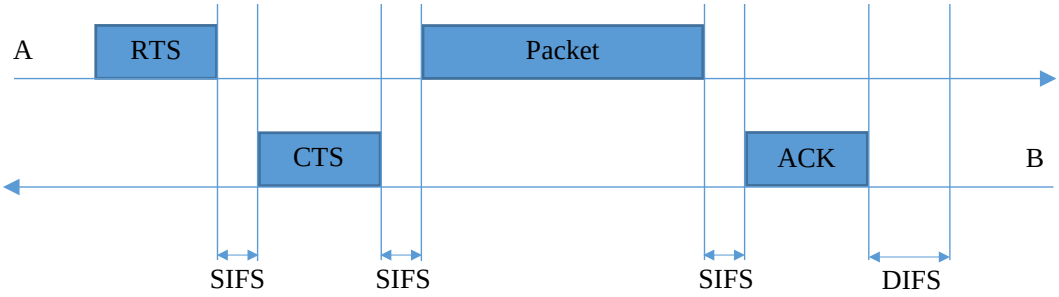


FIGURE 3.1. A successful CSMA/CA MAC protocol transmission along the link connecting A to B. All stages of the four-way handshake and mechanism illustrated must be completed for the transmission to be successful. Time proceeds to the right. The axes' arrows indicate the direction of each stage of the transmission between A and B.

C1. If A hears no-one (for a time period equivalent to what is know as one DIFS), then

S2. A says to B: "Can I send you a message?". This communication is known as the RTS (Request To Send).

C2. If B hears A's request, and is available to receive a message then

S3. B says to A: "Yes, send me a message". This communication is known as the CTS (Clear To Send).

C3. If A hears B's reply, then

S4. A says to B: "Lorem ipsum …", i.e., he sends the message, known as the *packet*.

C4. The protocol is self-checking for errors in that B will be aware at the end of the transmission if something has gone wrong. If nothing goes wrong and the message is received in full without errors, then

S5. B says to A: "I got it.". This is known as the ACK (Acknowledgment).

C5. If A receives the acknowledgement then he knows that the message has been successfully transmitted; the cycle ends, and he may consider sending further messages.

The communication may fail at any one of the conditional steps C1–C5. In this case, A waits for a given time before returning to step S1. Depending on the exact version of the protocol, this *backoff* time has a random element and increases exponentially over a bounded range if there are successive failures to send the same message (see Fig. 3.3 for more details of the backoff mechanism).

It is instructive to consider the mechanisms for how each of the conditional steps C1–C5 might fail. Firstly, C1 fails quite naturally if

M1. there is already a conversation within earshot of A.

Condition C2 fails, if either

M2. A speaks so softly that B cannot hear (or equivalently B is far from A);

M3. if there is a second ongoing conversation that B can hear but A cannot, which drowns out A's message; or alternatively

M4. a third person senses quiet and starts speaking at the same time as A.

M5. B fails to respond to A within a pre-defined time window (CTS Timeout).

Condition C3 fails similarly to the 2nd and 3rd mechanisms above, if

M6. A speaks too softly that B cannot hear all the message;

M7. a third person, having failed to hear A and B's previous correspondence, senses quiet and starts speaking at the same time as A.

Condition C4 fails, never generating the ACK if mechanisms M6 or M7 prevent the packet from successfully transmitting.

Condition C5 fails similarly to mechanisms M2 and M3 and further to M6 and M7, if

M8. B speaks too softly that A does not hear the ACK;

M9. a third person, having failed to hear A and B's previous correspondence, senses quiet and starts speaking at the same time as B transmits the ACK.

M10. the ACK is not received by A within a pre-defined time window (ACK Timeout).

If condition C5 fails, the ACK is not received and A assumes B has not received the message.

If the transmission is successful and a second packet is queued, A must wait for a period, equivalent to one DCF Interframe Space (DIFS) then employ a random backoff before it can attempt to transmit again, re-starting the process from step S1.

A Short Interframe Space (SIFS), is the period of time required for a device to process a received frame and to respond with a response frame. The SIFS is the processing time required between each stage of the conversation between A and B. It is the difference in time between the RTS being received and the CTS being transmitted, or further the equivalent time between CTS and Packet or Packet and ACK. Fig. 3.1 provides an illustration of a successful conversation between A and B.

If the transmission fails at any step (S1-S5), from any condition (C1-C5) as a result of any mechanism (M1-M10), the transmitter will backoff before re-attempting to transmit the same

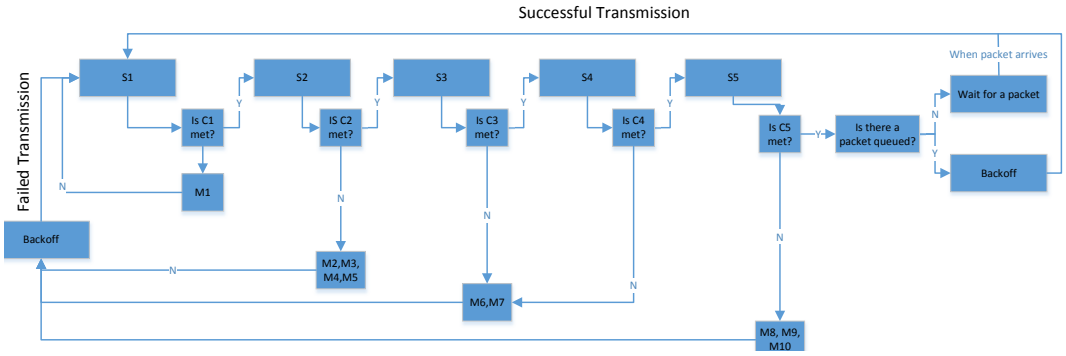


FIGURE 3.2. CSMA/CA flow diagram illustrating the relationship between the five steps of the CSMA/CA MAC protocol (S1-S5), the conditions that determine if each step is successfully completed (C1-C5) and the mechanisms by which each step may fail (M1-M10).

packet, starting the process over again from the beginning. The backoff time window increases with each re-transmission attempt. Fig. 3.2 illustrates the relationship between the five steps of the CSMA/CA MAC protocol, the conditions that determine if each step is successfully completed, and the mechanisms by which each step may fail.

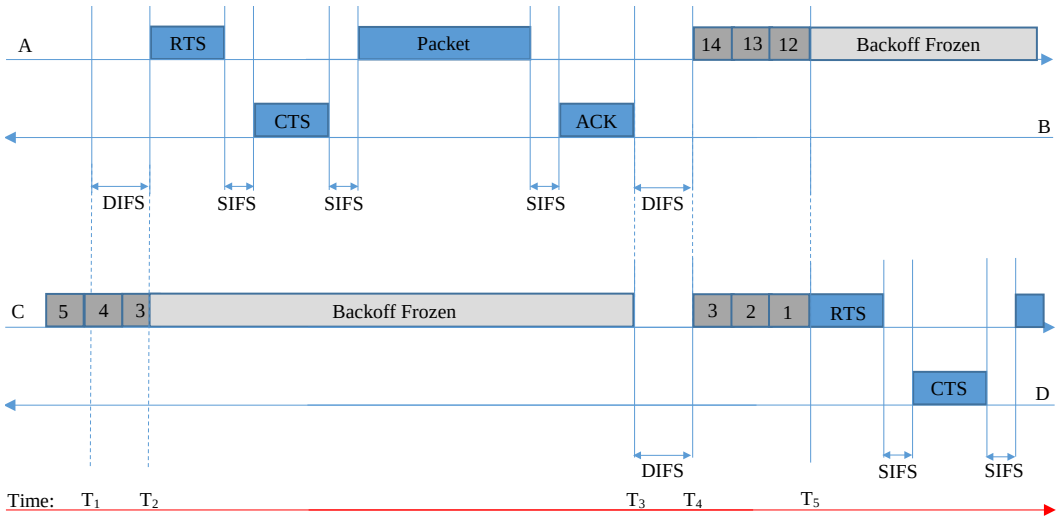
3.1.2 Managing Competing Demands in CSMA/CA Wireless Networks

The backoff mechanism is a key stage of the MAC protocol enabling multiple stations to fairly share the same wireless medium without a central controller. DCF adopts the slotted binary exponential backoff algorithm. A station desiring to initiate transfer of data shall use both the physical and virtual carrier sense functions to determine the state of the medium. A station with a new packet to transmit monitors the channel activity. If the channel is idle for a period of time equal to a DIFS, the station transmits. If the medium is busy the station shall defer: after the DIFS time, the station shall generate a random backoff period, after which the channel will be sensed again. This process reduces the probability of collisions during contention [189]. In addition, to avoid one transmitter unfairly dominating the channel (channel capture), a station must wait a random backoff time between two consecutive packet transmissions, even if the medium is sensed idle during the DIFS time after the ACK [29].

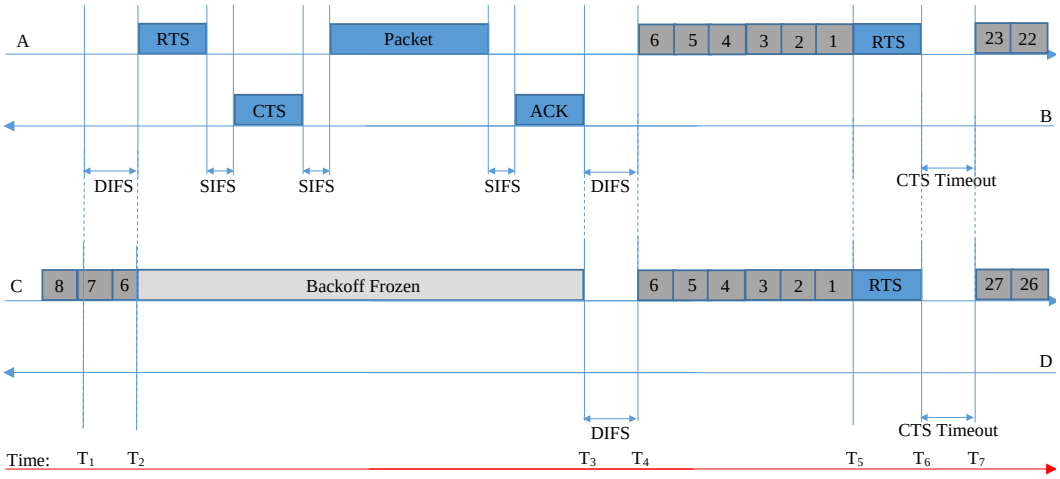
The backoff time, that is, time that must be deferred before the start of a transmission, is calculated each time a transmitter must backoff by

$$T_b = \mathbb{Z}(WU(0,1))T_s \tag{3.1}$$

where T_b is the backoff-time and T_s is the slot time. The product of $U(0,1)$, a pseudo-random number drawn from the uniform distribution 0 to 1, and W, which is an integer between $W_{\min}$ and $W_{\max}$ inclusive, must be an integer as indicated by $\mathbb{Z}$.



(a) Succesful contention resolution.



(b) A collision.

FIGURE 3.3. The backoff mechanism. (a) & (b) both consider a pair of links connecting A & B, C & D and the initial condition; A has nothing to transmit, C is in backoff having just transmitted a packet with other packets queued. For (a) At time T_1: A receives several packets in its queue that it wants to send to B; it must first sense the channel as quiet for a DIFS period. T_2: A senses the channel quiet and starts transmitting, C senses A's transmission and freezes its backoff counter. T_3: A finishes its transmission, both links must wait for a DIFS period. T_4: A generates a random backoff from a defined uniform distribution, C resumes the backoff process from the beginning of the interrupted slot. T_5: C's backoff counter reaches zero and begins transmitting, A freezes its backoff counter. For (b) T_1 to T_4 are as (a). T_5: C's backoff counter reaches zero and begins transmitting, simultaneously the backoff counter of A reaches zero and begins transmitting. T_6: A and C both listen for a CTS. T_7: Neither A or C successfully receive a CTS within a set time (CTS Timeout), as the simultaneously transmitted RTSs could not be interpreted by B or D. Hence A and C backoff a second time now selecting a backoff time from a larger range (double the previous).

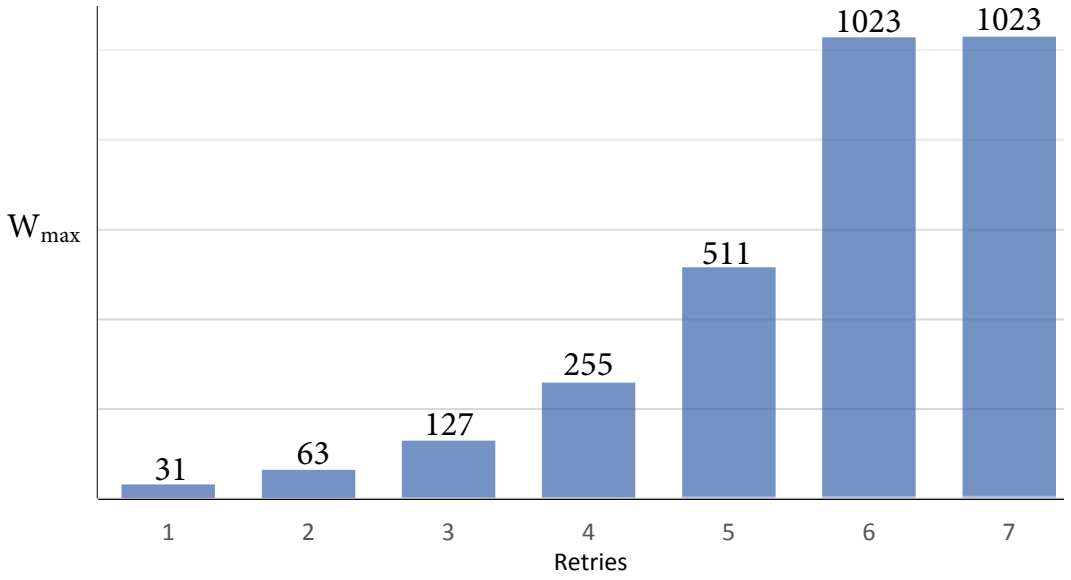


FIGURE 3.4. Backoff retries sequence for 802.11. The figure illustrates the exponential growth in the backoff window from W_{min} to W_{max}. As defined in the IEEE 802.11 standard [119], here from $W_{min} = 31$ to $W_{max} = 1023$

Each time a transmitter begins to backoff, a random integer value is selected between zero and W. The W parameter takes the initial value of W_{min} and its value is doubled, plus one, every time an attempted transmission is unsuccessful until reaching the value of W_{max}. It then remains at the value of W_{max} until it is reset (see Fig. 3.4). This improves the system stability in the case of high traffic demand. After a successful transmission, W is reset to W_{min}. The values of W_{min} and W_{max} are dependent on the particular standard.

The backoff is invoked by 802.11 CSMA/CA whenever a station wishes to transmit a packet and senses the medium busy. A station performing the backoff procedure monitors the medium for physical activity during each backoff slot. If no physical medium activity is sensed continuously for the duration of a slot, then the backoff procedure decrements its backoff time by one slot. If at any time during a slot there is physical medium activity detected, then the backoff procedure is suspended, that is, the backoff timer does not decrement in that slot. The medium must be sensed as idle for the duration of one DIFS time, before the protocol allows the backoff procedure to resume, from the beginning of the interrupted slot. Transmission only commences when the backoff timer decrements to zero. The result of this process is that when multiple stations are deferring and enter random backoff, then the station selecting the shortest backoff-time using the random function will win the contention [189] (see Fig. 3.3(a)). If two contending stations' backoff timers reach zero at the same time, both will generate RTS's which will result in a collision, triggering another random backoff interval for each selected from a new larger range (see Fig. 3.3(b)).

Seminal work by Bianchi [29] provided a performance analysis of the MAC protocol. Through-

Table 3.1: CSMA/CA MAC protocol parameters from [29].

Channel bit rate	1 Mbit/s
Packet payload	8184 bits
MAC header	272 bits
PHY header	128 bits
ACK	112 bits + PHY header
RTS	160 bits + PHY header
CTS	112 bits + PHY header
SIFS	28 μs
DIFS	128 μs
CTS timeout	300μs
ACK timeout	300μs
Slot time	50μs
$W_{\min}$	16 slots
$W_{\max}$	1024 slots

out his analysis, he uses the parameters displayed in Tab. 3.1. From observation of the MAC protocol, under the parameters derived by Bianchi in [29], it is possible to observe a ratio of time spent carrier sensing to actual data transmission of approximately 14% to 86% respectively. To simplify matters, Bianchi uses a channel bit rate of 1 Mbit/s, therefore the time to transmit 1 bit is equal to $1\mu s$. The backoff is equal to a randomly selected integer number of slot times between $W_{\min}$ and $W_{\max}$ where one slot time is $50\mu s$. More recent protocols may use higher bit rates and shorter times for individual protocol components, however, generally speaking, the carrier sensing components of the protocol (RTS, CTS, ACK) are always short in comparison to the packet transmitted. Bianchis' work demonstrates that the CSMA/CA protocol can reduce collisions and improves performance in networks with high demand and high numbers of nodes, however, the four-way handshaking process involves a significant overhead compared to a basic access scheme.

3.1.3 Applications of and Issues with the CSMA/CA Protocol

A conversation between A and B can only occur if both are positioned such that the 'volume' of the transmission allows each to be heard by the other over any background interference, expressed formally as a power ratio in decibels (dBm). Of course, the 'volume' at which A can transmit is independent of B and the two are not necessarily the same.

Further, a conversation between A and B is only possible if both consider the surrounding environment 'quiet' enough (sufficiently low interference). This again is an independent characteristic of each based on their CCA threshold (units dBm), a mechanism for determining whether the medium is idle or not. A low CCA threshold would result in a node more sensitive to interference, hence being less likely to consider the medium free. A higher CCA threshold would allow a higher level of interference before a node considers the channel busy. For A to hear B, received power

above a CCA threshold value must be maintained. If the CCA threshold of A were infinitely high, an infinitely high TP would be required for B to be heard, conversely, CCA threshold (for A) close to zero would mean transmission from B at very low volume could be heard.

Formally, node 1 is able to transmit to, or interfere with, another node 2 if

$$\|\mathbf{x}_1 - \mathbf{x}_2\|^2 < \mathrm{TP}_1 \mathrm{CCA}_2 \tag{3.2}$$

where $\mathbf{x}_1$ and $\mathbf{x}_2$ are position vectors of nodes 1 and 2 and TP_1 and CCA_2 their respective TP and CCA threshold. Their connection is dependent on the power density at the receiver which scales with transmit power TP_1 and the inverse square of the separation $\|\mathbf{x}_1 - \mathbf{x}_2\|$ of the transmitter and receiver. Note further, that by thinking of CCA sensitivity in terms of connectivity distance, it should scale reciprocally with CCA measured as a power density threshold.

Physical channel conditions may vary significantly from one location to another with background interference generated from other sources e.g., microwaves and bluetooth. Further interference and signal propagation delay / decay may differ depending on the environment (e.g., urban or rural) and objects / material (e.g., concrete buildings) between or surrounding the nodes slowing transmissions or reflecting generating self-interference. Over just a small distance in an urban area the environment can change significantly and thus the quality of a link may vary.

Independent TPs, CCA thresholds, and varied channel conditions result in asymmetric connectivity and interference patterns across networks. Instantaneous powers at the receiver for multiple transmitters will generally not be the same due to the different propagation decays reflecting channel conditions. Therefore, if more than one packet is transmitted to a destination at the same time, rather than both failing one may still be successful. For example, user A and user B positioned equal distance from each other and a third user C, sense the channel quiet simultaneously and attempt to transmit to C. The received power at C from A and B will likely differ. If this difference is sufficiently large one (A say) may be successful in its transmission where the other will fail. Once the packet transmission from A to C has started then only a level of interference sufficiently high such that C can no longer receive, will cause the packet to fail (capture effect [233]).

In the simplest analyses that one can develop, one supposes in effect that there are groups of individuals who share a small room so that every person can hear every other person loud and clear without delay. Hence only one person can speak at once, else messages become garbled. This is known as idealised MAC and considers a perfect physical layer. In particular, the described mechanisms M2, M3, M5, M6, M7 and M8 (p.49) do not occur.

Under these conditions, a transmitter will not attempt a transmission if another transmission is already occurring (M1). The MAC protocol can only fail as a result of two transmitters generating RTS at the exact same moment (M4). If two RTSs are generated simultaneously, they will always collide.

The MAC protocol has been designed to handle much more general set-ups, where (in effect) not all persons can hear each other. The signal to interference plus noise ratio (SINR) model has

been widely used to capture the impact of interference on network connectivity [76, 107, 108]. The SINR is defined by

$$\text{SINR} = \frac{S}{I+N}, \tag{3.3}$$

where S is the power of the incoming signal of interest, I is the interference power of the other (interfering) signals in the network and N is a noise term. Let $\mathbf{x}_1$ be the position vector of node 1. A node 2, position vector $\mathbf{x}_2$, is capable of receiving the transmitted signal from a node 1 (i.e., there is a direct link between 1 and 2) if the SINR is above a threshold β, i.e.,

$$\text{SINR}(\mathbf{x}_1 \rightarrow \mathbf{x}_2) := \frac{S(\mathbf{x}_1, \mathbf{x}_2)}{I+N} \geq \beta, \tag{3.4}$$

where N is the background noise and I is interference from other nodes in the associated area transmitting at the same time as node i thus causing interference. The function $S(\mathbf{x}_1, \mathbf{x}_2)$ defines power attenuation between $\mathbf{x}_1$ and $\mathbf{x}_2$ [275] and hence gives the received signal strength of the transmission from node 1 at node 2. Radio and antenna engineers typically use the following simple Friis Transmission Formula [88] for the signal power attenuation between two isotropic antennas in free space,

$$L = 20 \log_{10}\left(\frac{4\pi d}{\lambda}\right), \tag{3.5}$$

where L is the path loss in decibels, λ is the wavelength and d is the transmitter-receiver distance (i.e., $d = \|\mathbf{x}_1 - \mathbf{x}_2\|$) in the same units as the wavelength.

Fig. 3.5 illustrates the effect of SINR on the ability of two receivers, $\mathscr{R}_1$ and $\mathscr{R}_2$, to receive from their associated transmitters, $\mathscr{T}_1$ and $\mathscr{T}_2$ respectively. When the two transmitters $\mathscr{T}_1$ and $\mathscr{T}_2$ are transmitting simultaneously, their signals generate interfere at the other's receiver (the one that is not the transmitter's intended destination). In this example, SINR_1 and SINR_2 must be above a certain threshold value for the transmission to be successful. From the diagram, at $\mathscr{R}_2$, SINR_2 is illustrated such that the received signal S_2 is significantly greater than interference I_2 whereas at $\mathscr{R}_1$, SINR_1 is illustrated such that S_1 is slightly less than I_1. The SINR threshold may be set such that, when both $\mathscr{T}_1$ and $\mathscr{T}_2$ are simultaneously transmitting, the transmission from $\mathscr{T}_2$ to $\mathscr{R}_2$ is successful but $\mathscr{T}_1$ to $\mathscr{R}_1$ fails.

The capacity of a communication network can be defined as the maximum possible data-rate (bits/sec) that can be reliably transmitted from the source and received by a sink [108]. The capacity is considered an upper bound that, in reality, can never be reached according to a traditional information theoretic perspective. Considering a single transmitter-receiver link with a bandwidth subject to Additive White Gaussian Noise (AWGN), the maximum capacity over the channel is given by the Shannon-Hartley formula [209]

$$C = B \log_2(1 + \text{SINR}), \tag{3.6}$$

where C is the channel capacity (bits/sec) and B is the bandwidth (Hz).

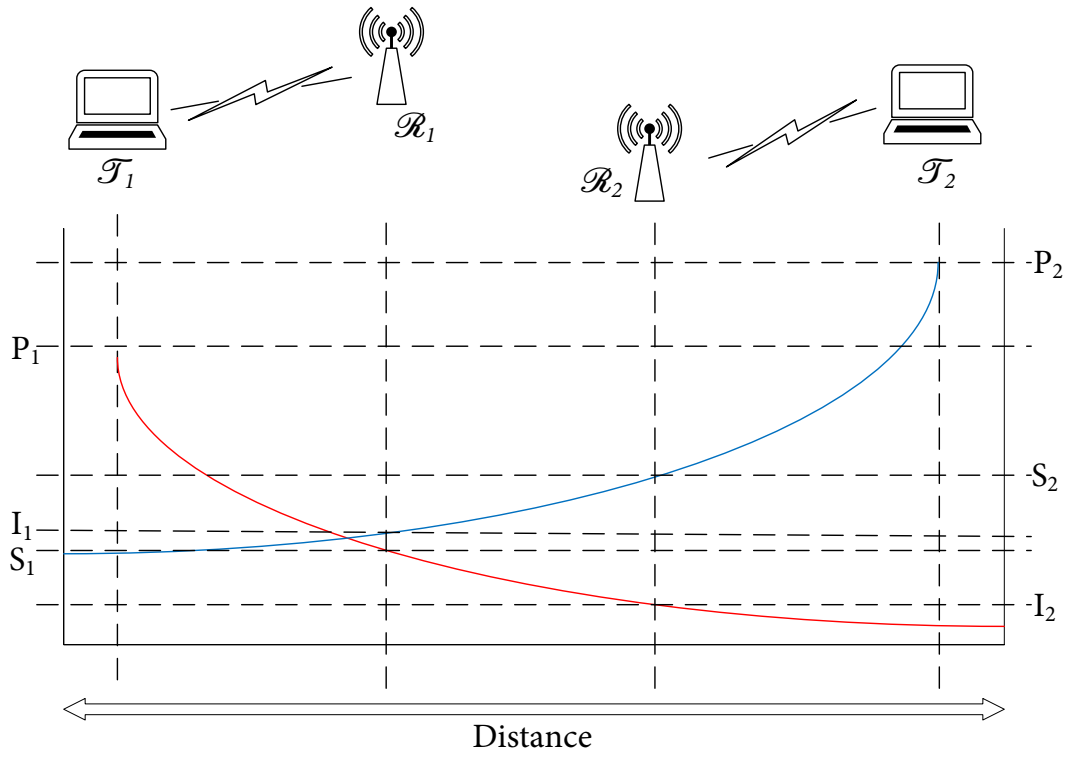


FIGURE 3.5. The effect of SINR on the ability of two receivers, $\mathscr{R}_1$ and $\mathscr{R}_2$, to receive from their associated transmitters, $\mathscr{T}_1$ and $\mathscr{T}_2$ respectively. Here, P_1 and P_2 indicate the TP at $\mathscr{T}_1$ and $\mathscr{T}_2$ respectively. Further S_1 and S_2 denote the received signal and I_1 and I_2 the received level of interference at $\mathscr{R}_1$ and $\mathscr{R}_2$ respectively. A constant level of background noise N_0 is assumed. The red line indicates the power of the signal transmitted by $\mathscr{T}_1$, dissipating over distance, to a received signal strength at $\mathscr{R}_1$ and resulting interference at $\mathscr{R}_2$. The blue line indicates the power of the signal transmitted by $\mathscr{T}_2$, dissipating over distance, to a received signal strength at $\mathscr{R}_2$, and resulting interference at $\mathscr{R}_1$.

Defining the capacity of a network becomes more difficult as the complexity of the network increases. Consider a network of n transmitter-receiver pairs. The demand is defined by an $n \times n$ matrix $\mathbf{D}$. The D_{ij} entry of that matrix represents the desired data-flow along the link connecting node i to node j. Each flow is affected by interference or noise from signals being generated by the remaining nodes transmitting in the network. These may have the effect of either interfering with the transmission between i and j, or alternatively, could aid communication. Due to the many possible interactions in a complex network, the capacity region, which is the set of all possible $\mathbf{D}$, is difficult to characterise [108]. Descriptive theories that consider the strict information-theoretic standard are needed [9]. (The capacity region is discussed in greater depth in Sec. 3.2.5 p.75).

It can be ascertained that the maximum rate at which a system can transmit information without errors, is limited by the bandwidth, the signal level, and the noise level. To relate SINR and bit error rate (BER), the number of bit errors per unit time, the energy per bit to noise power

spectral density ratio ($\mathscr{E}_b/N_0$) is typically used providing a normalised SINR measure (sometimes called SINR per bit). BER as a function of the $\mathscr{E}_b/N_0$ is then given by

$$\text{BER} = \frac{1}{2}\,\text{erfc}(\sqrt{\mathscr{E}_b/N_0}) \tag{3.7}$$

which indicates an inverse relation between SINR and BER. The BER is an indication of how often a packet has to be retransmitted because of an error. If SINR is such that the BER is too high, transmissions between a pair of nodes cannot be achieved.

Returning to the layman's description (Sec. 3.1.1) and considering the reality of the PHY medium, one may understand that the protocol could fail at any step (S1-S5), from any of the conditions (C1-C5) as a result of any of the mechanisms (M1-M10). The methods by which the protocol can fail at each step are summarised simply as follows:

- A listens (S1) to the channel and hears it's not clear, therefore waits for the channel to be clear.
- A asks B are you free to receive (S2); B does not hear the request and subsequently does not reply.
- B replies to A's question saying 'yes' (S3) but A does not hear the reply.
- Whilst A is transmitting (S4), another node nearby starts transmitting, thus B can hear transmissions from multiple nodes and cannot interpret A's.
- A does not hear an acknowledgement from B (S5).

Further, due to the reality of imperfect PHY and asymmetric interference patterns, the classic hidden node problem [126] (Fig. 3.6) occurs when a transmitter is audible from a receiver, but not from other transmitters communicating with that same receiver. This leads to difficulties in the MAC layer, managing multiple simultaneous transmission attempts to the same receiver only capable of receiving one signal at a time.

To understand the hidden node problem, consider a development of the layman's description of the MAC protocol (Sec. 3.1.1). A wants to talk to B. C also wants to talk to B. A and C are positioned on opposite sides of B such that they can both hear B but not each other. For a transmission to be successful, A to B say, B must notify the other transmitter, C, that it is busy and must wait.

The RTS/CTS mechanism employed by the 802.11 MAC protocol is designed to combat the hidden node problem. However, applying RTS/CTS in real networks does not always eliminate the hidden node problem. In [274], the authors argue that when the carrier sensing range is greater than twice the transmission range, RTS/CTS is no longer needed, rather, basic access (i.e., simple listen and wait) will perform as effectively.

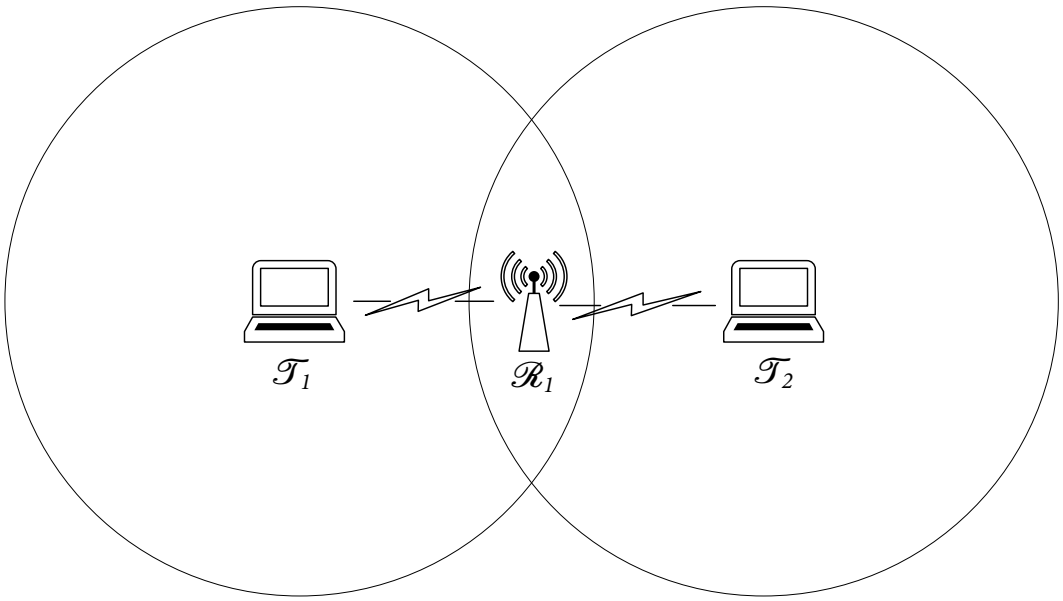


FIGURE 3.6. Hidden node problem. Transmitters $\mathscr{T}_1$ and $\mathscr{T}_2$ both wish to transmit to receiver $\mathscr{R}_1$. $\mathscr{R}_1$ can hear both $\mathscr{T}_1$ and $\mathscr{T}_2$ but, $\mathscr{T}_1$ and $\mathscr{T}_2$ are unable to hear each other. Hence there is a risk that $\mathscr{T}_1$ and $\mathscr{T}_2$ attempt to transmit to $\mathscr{R}_1$ simultaneously, resulting in a collision.

The exposed node problem [127] (Fig. 3.7), a further classic problem in wireless networking, occurs when a transmitter is prevented from sending packets because of a neighbouring transmitter. Consider a further development of the layman's example: suppose there are two transmitter-receiver pairs A, B and C, D where the two receivers (B, D) are out of range of each other, yet the two transmitters (A, C) are within carrier sensing range of each other. In this scenario, if a conversation between A and B is ongoing, C is prevented from speaking to D as it concludes after the carrier sensing step that its conversation will interfere with the conversation of its neighbours A and B. However D would be capable of hearing the transmission of C without interference because it is out of range of A. Exposed nodes prevent collision-free simultaneous transmissions and thus reduce spectral reuse.

The detrimental effects of hidden and exposed nodes reduce wireless CSMA/CA networks' performance. Numerous studies have examined the impact of hidden and exposed nodes (such as [122, 126, 127, 203, 280]) and make suggestions on how the impacts could be minimised. Often these improvements are limited to specific use cases / topologies. There is a great opportunity for more robust general solutions to provide a significant performance improvement.

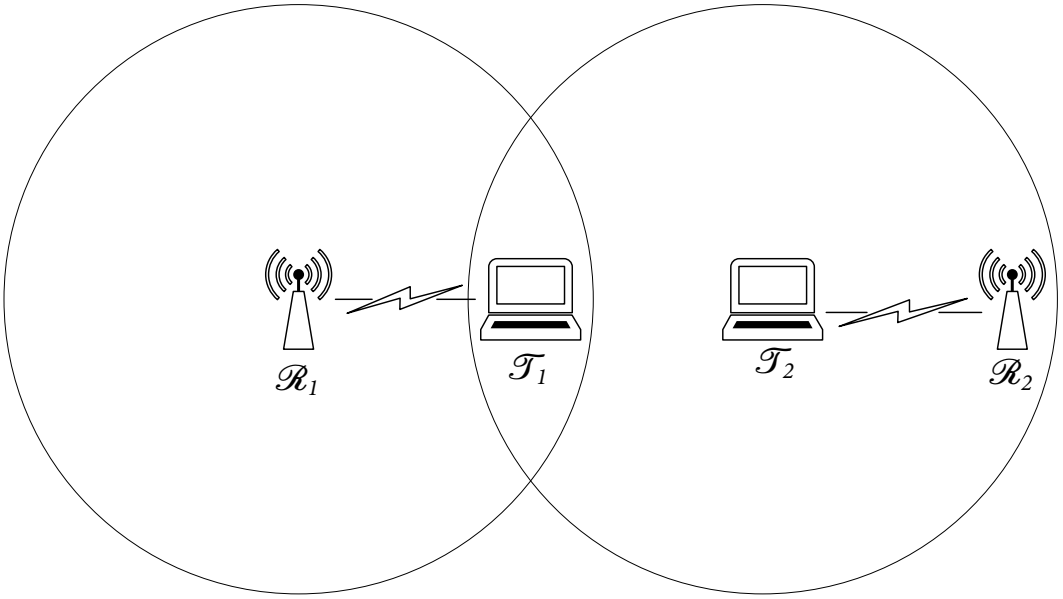


FIGURE 3.7. Exposed node problem. Transmitter $\mathscr{T}_1$ wishes to transmit to receiver $\mathscr{R}_1$ and $\mathscr{T}_2$ wishes to transmit to receiver $\mathscr{R}_2$. Because $\mathscr{T}_1$ and $\mathscr{T}_2$ are able to carrier sense each other's transmissions, they must follow the CSMA/CA listen and wait procedure (see Sec. 3.1.1). This prevents $\mathscr{T}_1$ and $\mathscr{T}_2$ transmitting simultaneously, although their respective receivers, $\mathscr{R}_1$ and $\mathscr{R}_2$, would be capable of receiving the signals if they did.

3.2 Review of Modelling and Simulation of Wireless Networks

In line with the definition of Whitworth [270], wireless communication systems are complex socio-technical systems and simulating them is a complicated but necessary challenge. The past decade has seen an exponential rise in data traffic, leading urban areas to require more access points, resulting in denser networks [60]: a trend expected to continue into an era of 5G and 802.11ax [10, 72]. Improved understanding through modelling and simulation studies will be essential in realising many of the technologies identified as key to the next generation of wireless technology in Chapter 2 (see Fig. 2.7 p.28 and Tab. 2.2 p.43) and [132]. The vast scale, high node density of modern networks and the unpredictable habits of their users present challenges in modelling them.

Within the systems engineering context, there is a need for models that aid in delivering complex technical systems by providing a quantifiable evaluation. Analytical mathematical models use mathematical techniques (e.g., symbols, logic, equations and graphs), where at least one of the variables is evaluative, to represent the system being modelled. The variables, typically measurable outputs (e.g., frequency of occurrence, volume of traffic, signal strength), are representative of system properties. The model describes the relations between these properties by a set of functions. Representing a system analytically in this way presents the worldview

that the sets of equations and probability distributions that comprise the model appropriately capture the characteristics of the system and further the output data, again in the form of a probability distribution, provides an appropriate macroscopic level of detail and information for understanding and developing the system. Analytical modelling is the typical mathematicians' approach to modelling a system and contrasts with the approach and worldview of normative simulation, the typical engineers' approach to modelling a system.

Normative simulation models aim to produce a faithful representation of the system for testing. Many systems are interconnected and subject to variability and complexity, both combinatorial and dynamic (i.e., relating to the number of components in a system and their relationships, and the interactions of the components over time respectively). Understanding (and predicting) the performance of systems that are subject to either variability, interconnectedness or complexity is difficult. Hence, normative simulation presents the worldview that predicting the performance of systems subject to all three is very difficult, if not impossible using purely analytical techniques. Simulation models, however, provide a tool that can be used to represent the variability, interconnectedness and complexity of any system and hence, it is possible for a simulation to generate some insight into system performance [215].

Normative modelling provides analysts with a method of testing a system design before committing to implementing it for real [232]. This methodology can produce very precise results but it can be difficult to apply the technique to large topologies with a high number of nodes and varying parameters. Models quickly become extremely complex and, generally speaking, even those attempting to maintain a high degree of fidelity in replicating the original system make simplifications. The computational complexity is a key constraint on this approach: models are difficult to build and can require long simulation run times.

A queuing system consists of packets moving through a network of nodes, forming queues at the nodes. Complex communications networks will consist of varying rates of arrival of packets and packets requiring services by one or many of the nodes [27], forming queues at nodes as they wait to gain access to the limited resource of bandwidth or channel access.

The typical mathematical approach to modelling networks of queues involves analytical queuing theory [63]. The main intention of analytical queuing theory is to establish the fundamental performance measures of the system which are probabilistic properties of the randomised variables [175, 238]. The interaction between a number of input variables (i.e., arrival rate and number of servers) are defined such that dependent outputs can be evaluated (i.e., service rate).

In contrast to analytical queuing theory, the traditional engineers' approach to modelling networks of queues involves normative simulation, typically normative discrete event simulation [86, 215, 229]. A discrete event simulator models a network which packets move through. Statistics in relation to individual packets or links within the network can be collected at a microscopic level. This may be extremely expensive in terms of CPU time. Key to any discrete event simulation is a simulation clock tracking the passage of time through the simulation, where

activity is not necessarily continuous. As simulation time progresses packets arrive in the system and compete for limited resources (i.e., channel access). They are served based on a set of rules defined in the simulator. Where packets cannot be served sufficiently quickly to match the rate of arrival, queues may form. At any stage, component parts of the simulator must be in a 'state', as defined by a set of programmed rules.

Analytical modelling and normative simulation are the two dominant approaches to modelling wireless CSMA/CA networks present in the existing literature. The remainder of this section discusses existing modelling and simulation approaches in relation to wireless networks. The discussion focuses particularly on CSMA/CA but also refers to studies investigating other protocols where useful ideas are raised. The section proceeds as follows: Sec. 3.2.1 discusses analytic queuing theory studies. Sec. 3.2.2 discusses normative simulation methods. Sec. 3.2.3 explores methods of modelling network connectivity. Sec. 3.2.4 explores how simulation methods capture the random characteristics of CSMA/CA. Sec. 3.2.5 looks at simulation performance metrics generated from analytical modelling and normative simulation techniques.

3.2.1 Review of Analytical Modelling of Wireless Networks

Seminal work by Bianchi [29] provided an analytical queuing theory model used to compute the 802.11 DCF throughput (i.e., rate of successful message delivery measured in bits per second (bit/s or bps)), assuming a finite number of nodes and perfect channel conditions (i.e., no hidden terminals or capture [117]). The work demonstrates that the performance of the model strongly depends on the system parameters, mainly the minimum contention window and the number of stations. Notably, he discovers performance to be only marginally dependent on the system parameters when RTS/CTS is enabled. The approach of the paper is in two parts. First, the behaviour of a single station is studied with the aid of a Markov model [175, 238], and the stationary probability τ that the station transmits a packet in a generic (i.e., randomly chosen) slot time is calculated. This probability does not depend on the access mechanism (i.e., Basic or RTS/CTS) employed. Further, by examining the events that can occur within a generic slot time, the author is able to express the throughput of both access methods as a function of the computed value τ.

Essential to the MAC protocol is the backoff process (more details of how the backoff is modelled can be found in Sec. 3.2.4). In Bianchi's model, a discrete timestep is adopted. The backoff counter is modelled as a stochastic process decrementing at the beginning of each timestep. The key approximation in the analytical model is that, at each transmission attempt, regardless of the number of retransmissions suffered, each packet collides with constant and independent probability ρ. It is intuitive that this assumption is more accurate as the contention window W and the number of transmitters n get larger. The probability, ρ is referred to as conditional collision probability, meaning that this is the probability of a collision seen by a packet being transmitted on the channel. Since the value of the backoff counter of each station is dependent

on its transmission history (e.g., how many retransmissions the head-of-line packet has suffered), the stochastic process is in reality non-Markovian. However, once independence is assumed, and ρ is considered to be a constant value, it is possible to model the bi-dimensional process with a discrete-time Markov chain.

Bianchi's method is demonstrated to accurately characterise the features of the protocol and their effect on the network throughput for a set of transmitters all within interference range of one another. The method can be used to identify the point at which transmitters, serving equal demands, become saturated and hence one measure of network capacity can be derived.

Gupta and Kumar [107], like Bianchi, use an analytic model considering a situation where all the transmitters in the network are required to transmit at the same bit-rate and further assumptions are made in relation to physical propagation. Their main result was the now common square root law which stated as the number of transmitters n increases, the per-transmitter throughput decreases as $1/\sqrt{n}$. Throughout their study, Gupta and Kumar assume perfect scheduling, i.e., no collisions. This is, of course, a simplification and hence their results provide a theoretical maximum achievable capacity, which it is fair to presume will be quite different from the actual achievable capacity.

Both the work of Bianchi, and of Gupta and Kumar, provide significant contributions but are limited in the detail they can provide. They are able to identify some performance characteristics e.g., the capacity of the system, however, are limited to the specific case of equal demand across all nodes. They are able to calculate the probability that any transmitter in the network will transmit successfully in a given time period but are unable to provide insight into the specific performance constraints on a single node in a complex topology.

Further examples of analytical investigations include work by Gao et al [91] which proposed a methodology to analytically compute throughput capacity or the maximum end-to-end throughput of a given flow in a multi-hop wireless network. The authors considered two key factors to affect the end-to-end throughput capacity; firstly neighbouring contentions, and secondly hidden node interference. Their contributions are made in proposing a contention graph to represent both neighbouring interference and hidden node interference. This considers neighbouring interference not only by the number of neighbouring nodes but also dependent on the relative location between neighbours. They propose a fixed point functional model for analysing the link capacity, and thereby the end-to-end throughput capacity of a flow in a multi-hop wireless network. To reduce the complexity the authors are forced to assume every transmitter is saturated, and compute the flow capacity as the minimum link capacity in the path.

Medepalli and Tobagi [172] present a model dependent on the computation of estimated service time for a single packet at the head of a transmission queue. Their analytical model can accommodate arbitrary topologies and enables some next-generation ideas to be captured such as directional antennas or multiple channels and further allows for different traffic demand per node. The model is used to investigated throughput starvation of some flows in a network. The

impact of the RTS/CTS process in preventing starvation is compared to modifying the contention window and found to have a less significant effect. Unfairness in IEEE 802.11 networks is further studied, identifying the trade-off between improved throughput and reduced fairness in multihop systems.

Work by Boorstyn et al. [33] modelled a wireless CSMA/CA network as a continuous time Markov process whose states are the independent link sets. The authors prove that the steady-state probabilities have a product-form solution (i.e., the metric for the collection of links can be written as a product of the metric across all the different links). Brazio [39] furthered this research showing that the product-form solution extends to a wider class of MAC protocols as long as the nodes' ability to carrier sense each other is symmetric. Papers [266] and [77] further utilised Boorstyns model to investigate the fairness problem in CSMA/CA networks, showing that unfairness is predominantly the result of topological inequalities, with nodes at the physical border of the network significantly favoured. Each of these authors uses a link-centric approach (as opposed to node-centric — see Fig. 3.8) to capture the behaviour of independent link sets (i.e., the links between two nodes). This approach enables the authors to capture slightly more sophisticated features of the CSMA/CA operation such as asymmetric interference patterns. However, these studies are limited by considering only saturated queues with exponentially distributed medium access attempts.

Further studies have extended the work of Boorstyn [33]. Garetto et al. [93], in contrast to Boorstyn, uses a node-based approach, modelling each flow in isolation; packet loss is considered as a function of the transmission probabilities and their interfering flows. Garetto favours this node-based approach, citing limitations in Boorstyns model such as its inability to capture the behaviour of the sophisticated random process in the 802.11 protocol. Specifically, it neither models packet losses due to the MAC protocol operation nor the Binary Exponential Backoff (BEB) mechanism. Another drawback is the demanding computation of all independent link sets in the network, which is an NP-complete problem. Garetto's approach enables the authors to derive the time duration that each node spends carrier sensing while attempting to transmit and provides insight into starvation problems. Although the authors provide some adaptations to improve the accuracy and level of detail generated from Boorstyn's model, they are forced to make several further simplifications of their own to manage the complexity. These include fixed transmission and carrier sensing range for all transmitters, no capture effect, and an error-free channel.

Unlike the majority of others, instead of focussing on fixed traffic demands, Jindal and Psounis's work [128–130] aims to characterise the capacity region (or achievable rate region), i.e., vectors of demand that can be met by the system, for a given topology. They do this by first finding the predicted service time of a specific link in terms of the collision probability of the link's receiver and its transmitter's perceived idle time. The complexity is to identify these collision probabilities and idle times as their values are dependent on the rates of other links in the system.

To simplify matters, the authors use a 'decompose-and-combine' approach. The local network topology is divided into multiple two-link topologies, and the values of the necessary variables for these two-link topologies are derived and subsequently combined. Calculating expected service time at all links enables the capacity region to be identified, which provides some insight into how neighbouring nodes affect each other's ability to meet demand. The analytical model assumes simple symmetric binary interference such that two incompatible nodes transmitting simultaneously will both fail. This interference model neglects some features of the 802.11 protocol like asymmetry, the capture effect [47] and multiple interferers [65]. The use of the capacity region as a metric for network performance is discussed in greater depth in Sec. 3.2.5 (p.75).

Relatively recently, Laufer and Kleinrock [151], and Kai and Zhang [135] provide interesting contributions that develop expressions for characterising the MAC layer by proposing analytical models. Their methodologies offer significant insight into the behaviour of wireless CSMA/CA networks and provide a method to predict their goodput (i.e., the application-level throughput or number of useful bits delivered per unit time excluding failed transmissions and packet overheads) performance. The papers recognise the challenge in creating a theory to characterise CSMA/CA networks stems from (i) firstly, the distributed nature of the protocol, which dictates that transmitters should back off from each other to avoid collisions; (ii) secondly the limited transmission range of nodes, which creates different broadcast domains whose behaviours are independent; and (iii) thirdly the buffer dynamics of unsaturated traffic sources, which sometimes causes queues to decrement to empty and alters the subset of nodes in contention for channel access.

To address these challenges Laufer and Kleinrock [151] consider a network of single hop flows communicating in a single radio channel with single transmission queues for each flow. The model captures buffer dynamics of unsaturated sources, while still respecting the interference constraints imposed by the wireless medium. The theory has no restrictions on the node placement, is suitable for arbitrary topologies, and provides a tool to characterise the capacity region assuming single hop flows and a single radio channel. An idealised CSMA/CA MAC protocol is assumed to control the medium access, as in several other works [33, 39, 77, 93, 161, 186, 266]. The model, like the CSMA/CA protocol, follows a listen and wait process; prior to transmitting a packet, each node first determines whether the medium is available via carrier sensing [143]. The model considers the channel to be busy if the received power exceeds a defined threshold; in this case, the node waits for the currently occurring transmission to end. Otherwise, the model recognises the medium is considered available and selects a backoff interval from a stochastic distribution, waiting at least this long before transmitting. The model defines separate transmitters on each node for different demands with a unique backoff counter recording the remaining time until transmission for each. If the channel becomes occupied by another transmission during the backoff interval, the node freezes its counter until the channel is idle again. When the counter

decrements to zero, the packet is transmitted (more detailed discussion on modelling the backoff can be found in Sec. 3.2.4).

The demands of each node in [151] follow an exogenous arrival process. The transmitters are not always saturated, i.e., sometimes there is no packet queued [33, 39, 77, 91, 161, 186, 191, 266]. Key to the methodology for characterising the capacity region in this paper is that demand to each node is controlled individually, unlike the works of Bianchi [29] and Gupta and Kumar [107] which assume demand for each node to be equal (for more discussion in relation to the capacity region see Sec. 3.2.5).

As in related literature [33, 77, 161, 259, 266] Laufer and Kleinrock [151] consider no interference when receiving packets, limiting the source of error to random noise and fading. This implies two things about the network being modelled. Firstly, hidden nodes are not considered, since the authors argued that their effect in unsaturated conditions is significantly reduced, and secondly, carrier sensing is instantaneous, i.e., a transmission is detected by those nodes in transmission range the moment it starts. The authors justify this assumption by arguing that carrier sensing only takes a few microseconds with modern wireless devices. Further to these assumptions, the probability of a successful transmission can be derived. For situations when transmission failures do occur, the model retransmits until the successful transmission is achieved, as is considered a good approximation of the CSMA/CA process [33, 161, 259, 266]. The model considers an area greater than the transmission range of any node, thus multiple transmissions can occur in parallel by nodes out of interference range.

Independently, Kai and Zhang [135] propose a mode similar to Laufer and Kleinrock [151]. Like Laufer and Kleinrock, but contrasting with most other works that consider saturated traffic, Kai and Zhang consider in their analysis the impacts of empty transmit buffers on the interactions and dependencies among links in the network. From the known starting point that an empty transmit buffer incurs extra waiting time for a link to compete for channel access, since when there is no packet waiting for transmission the link will not attempt to compete, the proposal of Kai and Zhang is that this extra waiting time can be mapped to an equivalent longer backoff countdown time for the unsaturated link. This yields a lower link demand intensity that can be defined by the mean packet transmission time divided by the mean backoff countdown time. This assumption is used by the authors to compute what they refer to as “equivalent access intensity” of an unsaturated link. This is a measure which incorporates the effects of the empty transmit buffer on the behaviour of channel competition, thus enabling goodput computation for CSMA networks. Specifically, they propose an iterative algorithm, to identify demands under which links are not saturated and use their “equivalent access intensities” to calculate link goodputs.

Kai and Zhang’s [135] algorithm is shown to be accurate under various offered load and protocol settings. The algorithm requires demanding computation in each iteration. For modest size CSMA networks, this is feasible. For larger networks (e.g., networks of more than 100 links), the computation could be extremely time-consuming.

3.2.2 Review of Normative Simulation of Wireless Networks

There are a number of simulation tools available to wireless researchers [54, 137, 226]. Network Simulator 3 (NS3) [196] and Optimised Network Engineering Tools (OPNET) [199], open source and commercial respectively, represent two of the most widely used simulators. Each offers comprehensive model libraries. As the telecoms industry has developed and become more complex, unsurprisingly the models required to represent it have become more complicated also. The current industry standard NS3 [196] had two simpler major versions precede it. Searches for 'NS3 simulator' in the Association for Computing Machinery (ACM) and IEEE digital libraries for the year 2015 yield 166 publications. In 2013, NS3 releases were downloaded approximately 24,000 times [196].

The simulators discussed in [54, 137, 226] represent a very different approach to modelling networks to the analytical approaches discussed in the previous section. Their object oriented design leads users to build a model from the level of individual components and links as opposed to a distribution of nodes with an assumed connectivity range. These methods are examples of normative modelling, capable of producing detailed simulation models and highly accurate representations of the real world.

When building models, even those attempting to maintain absolute fidelity to the original system make some simplification; typically they assume independent transmitters [172] and symmetrical interference patterns [91, 129]. A limitation of normative modelling is in the computational complexity [151]. Models are difficult to build and simulation run times are long. Further, the approach is limited by the data available, i.e., unless a complete system design is available many assumptions will have to be made and included in the model.

Normative simulation has great use in testing known protocols and topologies. For example, paper [211] uses NS3 to directly compare 802.11 releases 802.11n and 802.11ac and quantifies the performance improvement across a series of test cases. Paper [7] uses NS3 to compare the performance of Ad Hoc On-Demand Distance Vector (AODV), Destination-Sequenced Distance-Vector (DSDV), and Optimized Link State Routing Protocol (OLSR) for an application related to healthcare. Paper [59] further analyses the same routing protocols for application in VANETs (Vehicular Ad-hoc Network). Paper [283] uses OPNET to conduct a comparison of active queue management algorithms.

Many other studies use normative simulation tools to study new applications of wireless technologies: [162] investigates VANETs. Vehicles movements are restricted by road and environment and thus traditional random mobility models and way-point mobility model cannot reflect realistic vehicle traces. Hence, the paper investigates combining NS3 with SUMO (Simulation of Urban Mobility) [148] and uses SUMO vehicle traces as the mobility model in NS3 simulations. Paper [173] proposes a mobility model for an autonomous fleet of Unmanned Air Vehicles (UAVs) performing an area coverage mission. The UAVs, equipped with wireless ad-hoc capabilities, are required to achieve an optimal area coverage rate while maintaining connectivity with their

neighbouring UAVs and the base station. The proposed mobility model, in contrast to other models in the literature, considers energy consumption as a decision criterion. In order to assess the quality of the proposed mobility model, different performance metrics encompassing various quality measures have been used and tested in NS3. The analysis shows that the proposed mobility model outperforms existing models not only in terms of achieved global coverage and coverage fairness, but also in maintaining good network connectivity.

Some studies use normative simulation to understand performance limitations of certain protocols and wireless network features / ideas. Ng and Liew [191] used NS2 in their investigation into multi-hop ad-hoc networks. Their study focused on packet-loss, re-routing instability and unfairness problems at high demands. Their paper shows that controlling the offered load at the sources can eliminate some of these problems. To verify the accuracy of their NS2 simulation results, they further conduct a physical six-node multi-hop network experiment, confirming the accuracy of the simulator and reinforcing their finding that an optimal offered load can be found. The design of the experiment is simple with a manageable parameter space controlling the offered load to the one source node. Study [149] uses NS3 to test an algorithm for adapting the CCA threshold. The paper shows that CCA adaptation in this manner can provide a significant performance improvement, but due to the simulation complexity, the paper only considers static networks. Paper [249] evaluates the throughput performance of IEEE 802.11n WLANs using OPNET. It studies the effects of adjusting packet size, modulation and coding scheme, channel bonding, number of MIMO spatial streams, block acknowledgement and type of service / access category on maximum throughput and efficiency. To limit the parameter space the paper considers equal demands to all nodes and simple interference topologies.

Each of these studies provides interesting insight into ideas which could potentially realise significant network performance gains. However due to the complexity and computational demand of the simulations, they are forced to limit their study to a simple subset of the possible parameters, not necessarily representative of real-world networks. It is thus unclear from the conclusions whether the enhancements they suggest will realise the gains they show, when implemented in larger networks with different traffic demands or conditions.

3.2.3 Modelling Wireless Network Topologies

As wireless networks may come in many forms, the usual modelling approach is to generate random topologies. These topologies can be generated by defining positions of nodes in an area and providing their physical characteristics and hence connections (node-centric), or alternatively, a more abstracted view can define the characteristics of and relationships between links of the network (link-centric).

A Poisson Point Process (PPP) is used for representing apparently random spatial processes and is a common method of depicting network node position in a plane [110]. Other representations are adaptations of the PPP and numerous examples are discussed extensively in [110]. One

example is the Matérn Hard-Core Process which distribute a defined density of points under a PPP, then removes all points that fall within a defined distance of a neighbour. Effectively points are not allowed to be within a specified distance of each other. A Cox process is a development of a PPP which provides a method for representing models of clustered point patterns. It is generated by considering the intensity of a PPP as a random variable in itself. Hence, when first studied by Cox [62], he referred to it as doubly PPP. Cox's adaptation of the PPP leads to a distribution very different in appearance to the random distribution of the Matérn Hard-Core Process. Whereas the Matérn Hard-Core Process ensures no points are within a certain distance of each other, the Cox process provides points clustered together.

Typically analytical queuing theory methods imply this PPP distribution by defining a node density per area and hence a probability in relation to the node's ability to services its arrivals. Alternatively, a network may be prescribed explicitly, defining each node and its relationship to other nodes in the network.

Paper [53] investigated the distance that a node's broadcast penetrates a set of nodes randomly distributed according to a homogeneous PPP across an infinite two-dimensional plane. There are several other examples of a PPP approach being used in studies exploring characteristics of network topology [112, 193, 259]. For example, other studies favoured a Matérn Hard-Core Process [17, 41, 42], using it to model interference between sets of concurrent transmitters in a CSMA network. Problematically these Matérn Hard-Core Processes become difficult to analyse with increasing network complexity, and therefore researchers often favoured using a PPP to provide an approximation of the distribution of concurrent transmitters. Paper [109] quantifies the resultant difference between hard-core and PPP approaches and concludes that PPP is a good approximation.

Gupta and Kumar [106] discuss how wireless signals sharing the same medium interact with each other by studying the connectivity of uniformly distributed nodes in a circular area. Transmissions intended for one receiver cause interference at other receiver nodes. This impacts the SINR at the latter receivers, and hence, lowers their information processing capacity. Thus they recognise the need to control the transmitter power such that transmissions reach their intended receivers while causing minimal interference for other receivers sharing the same channel. They derive expressions determining the critical power at which each node needs to transmit so as to guarantee complete connectivity of the network, and hence the interference received by any node in the network from any other's transmission.

Contributions by analytical investigation of connectivity within bounded areas have been made in [24, 25, 74, 223, 224, 242]. Each uses a simple model to characterise the wireless channel, connecting nodes based on a fixed threshold Euclidean distance. This model resembles a special case of the SINR model simplifying the interference of additive signals to simple binary collisions between simultaneously transmitting nodes. Some authors argue these pure geometric models are limited in their ability to model a real network situation because of their deterministic

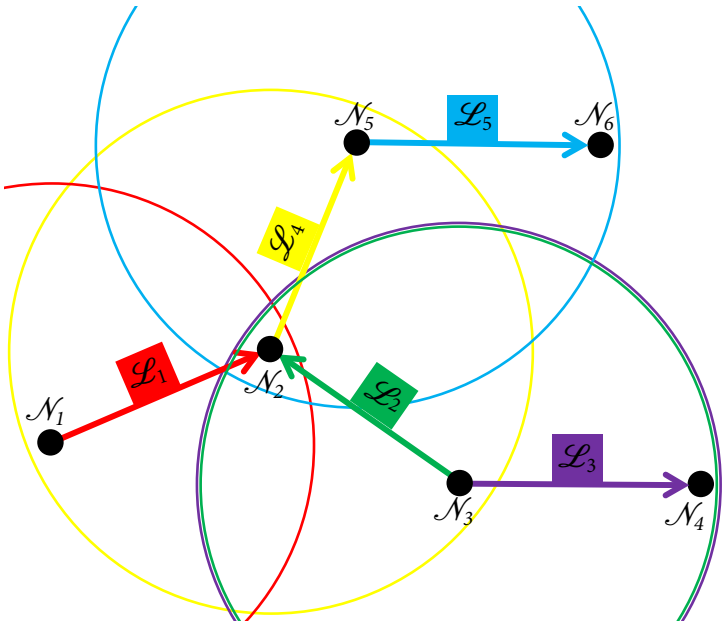
nature, and that for a more realistic representation, distance-dependent channel models must be developed. For example papers [284] and [239] showed that improved models of the network layer depend on more accurate modelling of the physical layer for wireless multihop networks. A significant challenge is that applying the SINR model (incorporating realistic interference) significantly increases the complexity of modelling these situations. The formation of a direct communication link between a pair of nodes is affected by both the locations and the concurrent activities of other nodes in the network. This is a considerable step in complexity from the fixed threshold model or the random connection model because connections can no longer be assumed to be independent.

Papers [113, 182, 183] investigate necessary conditions for networks to be connected under the more realistic log-normal connection model, i.e., two nodes are connected if the received power from the transmitter is greater than a given threshold, with the connection attenuation following the log-normal model [102]. These results, however, rely on the assumption that the node isolation events are independent, which is yet to be proven [275]. Considering the significant impact on the connectivity of interference resulting from concurrent transmissions, limited work exists on analysing connectivity under the SINR model. Papers [15] and [154] do study connectivity under SINR, however, the research is from the perspective of channel assignment. Channel / time slots are assigned to each link to enable the maximum number of links to be simultaneously transmitting while satisfying the SINR requirement, simplifying the model.

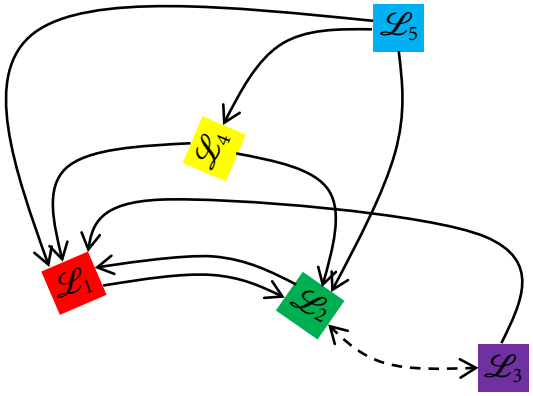
Paper [93] defined node or link-centric approaches to modelling networks. Node-centric models derive the individual throughput of each node via a function of the throughput of neighbours with which it shares the channel (e.g., [91, 129, 172, 191]). In link-centric models, the relationships between pairs of links and the behaviour of independent link sets are captured (e.g., [33, 39, 77, 151, 186, 266]). Link-centric models have advantages over node-centric as, despite still representing a simplified binary model, they are able to capture features of SINR that node-centric models are not, specifically asymmetric collision and channel access probabilities, where the relative position of nodes in a network can result in a high probability of collision at one receiver, but a high success rate in another. Fig. 3.8 illustrates the different views of node and link-centric models.

Normative simulators, such as NS3 [196], include MAC and PHY representation and hence include SINR interference models. Transmitted signal characteristics are defined by numerous propagation loss models available in the NS3 simulation library. These typically introduce probabilistic uncertainties which impact the MAC layer protocol.

Regardless of the approach chosen for modelling wireless network topology, an abstraction must be made and the potential limitations must be considered. Several of the discussed studies have provided significant contributions using simple binary interference models. The investigative philosophy is that a model is used as a tool for investigating a question of interest and is constructed accordingly.



(a)



(b)

FIGURE 3.8. Node and link-centric models. A network of six nodes ($\mathcal{N}_1, \mathcal{N}_2, ..., \mathcal{N}_6$) is connected by five uni-directional links ($\mathcal{L}_1, \mathcal{L}_2, ..., \mathcal{L}_5$). (a) A node-centric model / representation. Coloured circles represent the interference of each node transmitting along a corresponding colour link. The direction of the link indicates the direction of data flow. (b) A link-centric model / representation. Arrows indicate relationships between links. An arrow points to a link which will fail if it attempts to co-transmit with the link from which the arrow originates. The arrows between $\mathcal{L}_1$ and $\mathcal{L}_2$ indicate they will fail mutually if they attempt to co-transmit, because they have the same intended destination, which may only receive one signal at any time. If $\mathcal{L}_1$, $\mathcal{L}_2$ or $\mathcal{L}_4$ attempt to co-transmit with $\mathcal{L}_5$ they will fail but $\mathcal{L}_5$ will transmit successfully. This is because the transmission on $\mathcal{L}_5$ generates interference at the receivers of $\mathcal{L}_1$, $\mathcal{L}_2$ and $\mathcal{L}_4$, but they cannot be heard at the receiver of $\mathcal{L}_5$, hence the arrow only appears in one direction. Note $\mathcal{L}_1$ and $\mathcal{L}_2$ have no nodes in common with $\mathcal{L}_5$. The double-ended dotted arrow between $\mathcal{L}_2$ and $\mathcal{L}_3$ indicates they may not co-transmit as they share the same source node.

3.2.4 Modelling Random Backoff in Wireless CSMA/CA Networks

In a different setting to that described in Sec. 3.2.3, a Poisson Process or other process can be used to model events occurring seemingly at random in time. Within the context of analytical queuing theory, it is common to represent packet arrivals in this way. This process defines the arrival of each packet as an independent event not related to other arrivals. This is a typical method of modelling arrivals in many queuing systems where the arrival rate can be considered random, for example, phone calls arriving at an exchange. The method, however, may be less appropriate for other networks where arrivals are scheduled (such as buses arriving at a depot). An alternative may be to prescribe the arrival of events individually, but this, of course, requires some data in relation to the real system.

Analytical models use different types of distribution to represent aspects of the system, e.g., arrivals, services, failures, and hence derive system performance measures based on the relationships between these distributions. The normative approach to modelling would be to prescribe the relationships between component parts of the system and rules governing their interactions. Output, such as services and failures are then dependent on the arrivals (often still a random distribution) and further their interaction in the system.

The backoff is a key feature in determining the random nature of the CSMA/CA MAC protocol. Achievable throughput (and subsequently goodput) is largely determined by the backoff mechanism characteristics. However, different interpretations or implementations of the backoff mechanism exist in the CSMA/CA modelling literature: the standard is not a completely rigid specification.

Consider a link $\mathscr{L}_1$ between nodes $\mathscr{N}_1$ and $\mathscr{N}_2$. The RTS/CTS exchange is unsuccessful if either the RTS or the CTS is lost due to PHY errors or an RTS collision occurs at the receiver $\mathscr{N}_2$. Bianchi explains [29] that an RTS collision will occur only if the backoff counter of a second conflicting link ($\mathscr{L}_2$ say, connecting $\mathscr{N}_3$ and $\mathscr{N}_4$) also expires simultaneously. For efficiency reasons, DCF employs a discrete-time backoff scale. The time immediately following an idle DIFS is slotted, and a station is allowed to transmit only at the beginning of each slot time. An RTS collision will occur only if the backoff counter at the second link also expires in the same slot duration. The slot time size, σ, is set equal to the time needed at any station to detect the transmission of a packet from any other station (σ=50 μs in [29]).

As derived in [29], and further explained in [129], $P_{W_0}^{\mathscr{L}_2}$, the probability that $\mathscr{L}_2$'s backoff is equal to zero during the same time slot as $\mathscr{L}_1$ and hence a collision between RTSs will occur, is bounded from below by $2/(W_{\min}+1)$ and from above by $2/(W_{\max}+1)$, where $W_{\min}$ is the initial contention window and $W_{\max}$ the maximum contention window. Hence,

$$\frac{2}{(W_{\max}+1)} \leq P_{W_0}^{\mathscr{L}_2} \leq \frac{2}{(W_{\min}+1)} \tag{3.8}$$

defines the probability of a collision between RTSs (see Tab. 3.1 p.54 for the parameters of Bianchi's calculation).

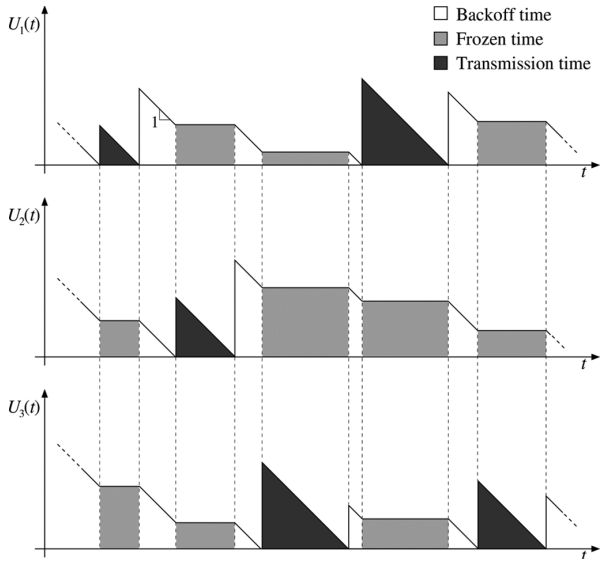


FIGURE 3.9. An illustration of Laufer and Kleinrock's model progressing with time for three saturated links within carrier sense range. Time proceeds to the right. The graphs show the unfinished work of each transmitter at any time, i.e., the remaining backoff or the remaining transmission time. The backoff of a link is frozen while the channel is in use by another. Figure reproduced from: R. Laufer and L. Kleinrock, "The Capacity of Wireless CSMA/CA Networks," IEEE/ACM Trans. Netw., vol. 24, no. 3, pp. 1518-1532, Jun. 2016.

The model applied in Laufer and Kleinrock's work [151] makes no assumptions about the backoff interval distribution. Several other works show it is not necessarily exponentially distributed [33, 39, 77, 93, 186, 266]. The [151] model defines separate transmitters on each node for different demands, with a unique backoff counter recording the remaining time until transmission for each. If the channel becomes occupied by another transmission during the backoff interval, the node freezes its counter until the channel is idle again. When the counter decrements to zero, the packet is transmitted.

Laufer and Kleinrock [151] consider backoff in continuous time, not slotted. Freezing of the backoff counter occurs the moment a contending station begins transmitting on the channel. The authors state among their assumptions that they neglect the possibility of collisions due to RTSs from multiple stations being generated simultaneously and that stations sense each other immediately, the moment one starts transmitting. Combined, these assumptions and this interpretation of the backoff mechanism produce a model that will achieve significantly better throughput than actual real-world CSMA/CA. Fig. 3.9 illustrates Laufer and Kleinrock's model progressing with time.

Through his analysis, Bianchi assumes the probability of collisions amongst RTSs at the lower bound of his identified range (effectively no collisions). In [129] Jindal and Psounis argue it would be more realistic to consider the upper bound. They conclude that this higher probability

of collision is more representative of the MAC protocol operating in a PHY environment with some propagation loss and fading.

Paper [272] discusses Bianchi's [29] use of Markov processes to analyse saturated 802.11 and comments that due to some erroneous transitions in the Markov chain used and some incorrect hypotheses, the model overestimates throughput in 802.11. Anderton and Young [8] present evidence that many commonly used models for designing contention resolution algorithms do not adequately account for the cost of collisions. Tinnirello, Bianchi, and Xiao [250] outline a number of issues and critical features raised by Bianchi's previously proposed model [29] and other related models. In particular, the authors focus on the backoff counter decrement rules. Different interpretations of how the backoff counter is frozen have led to different CSMA/CA models. The analytical model proposed in [29] implicitly relies on the assumption that the backoff counter is decremented at the beginning of a slot time. This implies that, at each slot time, the backoff counter is decremented, regardless of the fact that the slot is empty or contains a frame transmission (or collision). This conclusion is incorrect: looking again at the original 802.11 specification reveals that a more conforming modelling assumption is to decrement the backoff counter at the end of a slot time. In fact, the IEEE 802.11 standard [119] specifies in Sec. 9.2.5.2 that, if the medium is determined to be busy at any time during a backoff slot, then the backoff procedure is suspended. The authors conclude among other observations that, under saturation conditions, the slot immediately following a successful transmission can be accessed only by the station that has successfully transmitted in the previous channel access, by picking a new backoff counter equal to zero. Other stations with frozen backoff counters will resume counting down from the beginning of the integer backoff counter slot which had been frozen (see Fig. 3.3(a)).

Tinnirello, Bianchi, and Xiao [250] further identify that assumptions of infinite retransmission attempts greatly impact upon the probability of a collision, compared to a finite retransmission attempt limit. Assuming infinite retransmission attempts, the contention window grows from $W_{\min}$ to $W_{\max}$ and then remains at $W_{\max}$, whereas if a finite limit is reached the contention window will reset to $W_{\min}$, changing the probability of two backoff timers reaching zero simultaneously and thus leading to a collision.

The analytical models of Bianchi [29], Laufer and Kleinrock [151], Jindal and Psounis [129], Tinnirello, Bianchi, and Xiao [250] and others model attributes of the network performance, such as backoff characteristics and subsequently successful or failed transmissions, via probability distributions. Normative simulation tools such as NS3 [196], represent stages of the MAC protocol as a series of events and step through those events in a defined order. As the simulator reaches an event such as a backoff timer expiring, then a programmed decision is made as to which event to move to next. The backoff procedure is designed to faithfully represent the IEEE standard, hence independent objects within the simulation (i.e., nodes) separately select random backoff times, and a collision or successful transmission is dependent on the individual times the nodes' backoff counters expire.

Interestingly, study [28] draws the striking conclusion, that over a test of six commonly available commercial WiFi cards, not one performs exactly as expected in terms of backoff operation, despite all supposedly being designed to the same standard. Of the six tested cards, in some cases, implementation issues seem to impair the proper card operation. For other cards, the authors of [28] conclude, manufacturers deliberately alter the backoff parameters from the standard specification, in order to provide an unfair advantage for these cards with respect to the competitors.

3.2.5 Simulation Performance Metrics in Wireless CSMA/CA Networks

Typically analytical models generate macroscopic, i.e., flow-level probability distributions, whereas normative simulations generate microscopic metrics, i.e., individual packet queuing and service times. Both methodologies are capable of generating outputs describing packet services and queuing, and the level of detail required is a decision for the model's designer.

Statistics relating to individual demands can be useful, but it is often interesting to understand the interaction of multiple demands in competition for a limited resource. For example, consider a situation where a demand d_1 (packets per unit time) is being served whilst another node sharing the limited resource is serving a separate demand d_2. If d_1 increases by 10%, what demand can d_2 then achieve? Or further if an additional demand d_3 is introduced, what demand can d_1 and d_2 achieve maintaining stable network conditions?

One method of understanding this interaction is, rather than applying the same fixed demand to all transmitters, instead identify vectors of demand that can be met by the system. This may then enable a *capacity region* to be identified. Fig. 3.10 plots an illustration of two demands in contention. In this example, the capacity region boundary shows a concave shape, indicating a lack of efficiency when two demands are sharing the channel. Statistics provided in relation to the capacity region can enable improved understanding of potential network performance gains, compared to the simple measurement of saturated flows. The difficulty is that identifying the capacity region for high numbers of demand vectors is computationally demanding.

Most of the studies discussed so far in this section consider fixed demands and provide outputs using standard performance metrics such as throughput, goodput, latency and collision probability. The previously mentioned work of Jindal and Psounis [128–130] and further Laufer and Kleinrock [151] are among the few which aim to characterise the capacity region.

The results of Laufer and Kleinrock [151], when examining simple two and three node topologies, with all nodes within carrier sensing range (under perfect channel conditions assuming no delays or errors), show a convex capacity region. This implies that the introduction of more demands can increase goodput. This counter-intuitive result is unique in the wireless communications literature. Queuing theory applications in other areas, where the capacity region is a standard method of performance analysis (i.e., discrete traffic modelling), show the introduction of new demands always results in a performance penalty. The reason for the gain in [151] is

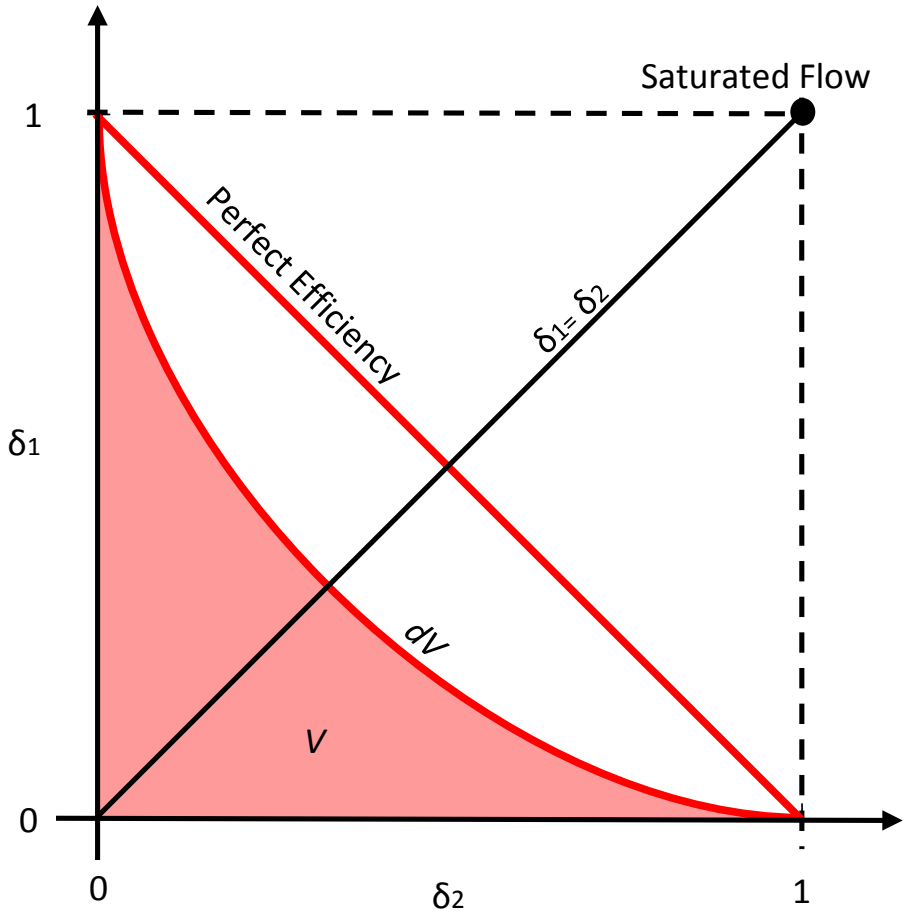


FIGURE 3.10. Schematic of a typical capacity region for two demands. The diagram illustrates the performance of a system when demands are in contention for a limited resource. Here δ is a non-dimensional demand intensity, with 1 indicating the maximum demand the system could serve for just one demand on its own (without contention). The shaded area, V (hear an area but generally n-dimensional), indicates demand combinations that are inside the capacity region. The curved line dV indicates the boundary of the capacity region. Its concavity indicates a loss in efficiency when two demands are in contention and a reduction in total goodput.

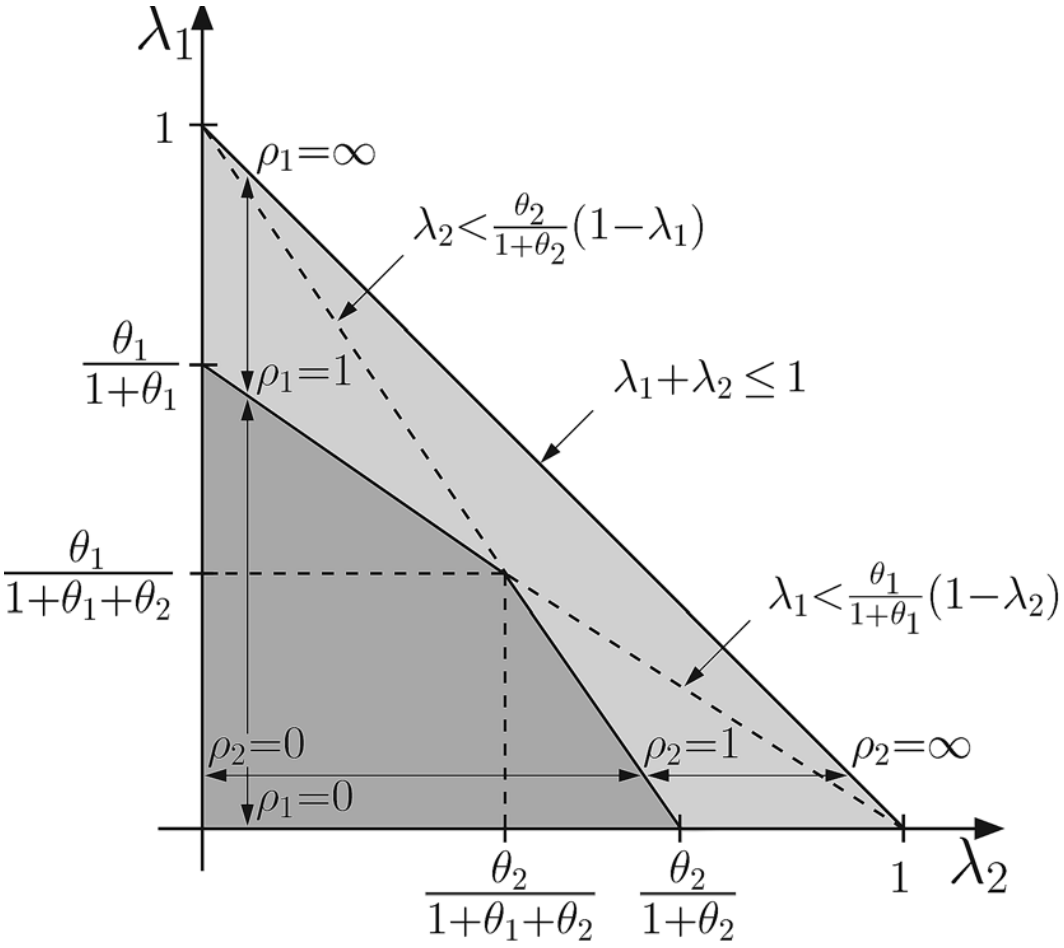


FIGURE 3.11. Laufer and Kleinrock [151] two-link convex capacity region indicating a gain in goodput when transmitters share the channel. In their notation λ is the input rate, θ a ratio between a random variable for transmission time and a random variable for the backoff interval. Further, ρ is a stability factor of the transmitter: it can be thought of as the utilisation factor and therefore as an indicator of how close to saturation a given transmitter is. Figure reproduced from: R. Laufer and L. Kleinrock, "The Capacity of Wireless CSMA/CA Networks," IEEE/ACM Trans. Netw., vol. 24, no. 3, pp. 1518-1532, Jun. 2016.

that at the end of a wireless CSMA/CA transmission, any node is forced to backoff for a random time period before attempting to transmit a further packet. This is to prevent one node from dominating the channel. This mechanism slows the maximum achievable transmission rate of one node transmitting alone, since with no competition for channel access, there are periods when no data is being transmitted. When multiple nodes are wanting to transmit, always the node that selects the shortest backoff period will transmit first. This has the effect of shortening the average time between data transmissions and hence increasing the capacity of the channel. (Further interpretations of the backoff were discussed in Sec. 3.2.4.)

Furthermore Jindal and Psounis [129] speculate that it might be possible to achieve a gain in goodput when adding additional demands and hence achieve a convex capacity region. Consider two links competing for channel access: the RTS/CTS exchange is unsuccessful if either the RTS or the CTS is lost (not heard by its intended receiver) due to physical layer losses or an RTS collision occurs. An RTS collision will occur only if the backoff counter at link two expires in the same slot as link one, resulting in both links sending an RTS packet. Bianchi [29] derives upper and lower bounds for the probability of such a collision occurring amongst RTS packets (see (3.8) p.72). Jindal and Psounis argue that approximating by the lower bound is accurate when assuming an ideal physical layer (i.e., there are no collisions and data losses); otherwise, approximating it with its upper bound will be more accurate. The impact of this change is significant. The small increase in collision probability results in a small increase in collisions but consequently, much more time is spent in the backoff and carrier sensing stages of the protocol, and hence much less time transmitting data. The slight convexity of the capacity region is lost, with the overheads of collisions in the imperfect protocol instead resulting in additional demands leading to a loss in goodput and thus a concave capacity region shape. The authors conclude the capacity region is concave under realistic physical layer conditions.

Jain et al. [123], Toumpis and Goldsmith [252], and Kodialam and Nandagopal [144] further explore the capacity region via analytical queuing theory methodologies. They use a convex combination of the capacities of the feasible link sets to derive network capacity regions. These works, however, assume a Time Division Multiple Access (TDMA) network with an optimal centralised scheduler and do not directly apply to CSMA/CA networks.

A few other papers use the capacity region to understand the interactions between links in CSMA/CA networks [19, 145, 236, 237]. Work by Subramanian and Leith [236, 237] presents an analytical characterisation of a non-convex capacity region for CSMA/CA WLANs discussing that, in addition to being of intrinsic interest as a performance metric, the capacity region can be exploited to facilitate optimisation approaches to be applied to CSMA/CA, especially resource allocation problems. Bai, Qiu and Xue [19] investigate fairness in IEEE 802.11, explaining that asymmetric knowledge of source or destination nodes leads to asymmetric capacity regions. Intrinsically deficient MAC scheduling implies key questions: firstly, how to define end-to-end fairness under IEEE 802.11 and secondly how to achieve it. The authors define a new fairness model for non-convex capacity regions.

Koseoglu and Karasan [145] state that one of the main drawbacks of the CSMA/CA protocol is the collisions that may occur as a result of the propagation delay between nodes. Generally, when modelling existing wireless configurations under ideal physical layer conditions, propagation delay is considered negligible in comparison to transmission times, and therefore not a problem. On the other hand, larger propagation delays should be considered in the performance modelling of future wireless networks. Wireless networks are now used to cover larger geographical areas, with obstructions in the line of sight, hence propagation delay is larger. The authors also discuss more

niche applications (e.g., underwater acoustic networks) which experience very large propagation delays due to low propagation speed. The results show how under realistic channel conditions, the capacity region of CSMA networks is concave and they demonstrate how that concavity increases with increasing propagation delay.

In summary, statistics provided in relation to the capacity region enable an improved understanding of potential network performance gains, compared to the simple measurement of saturated flows. The complication of using the capacity region as a metric is that for simulations of larger networks, an exhaustive search of the (very high dimensional) demand space will not be computationally tractable, and more refined search techniques will need to be adopted.

3.3 Identifying Opportunity for Innovation and Comparison Of Methodologies

Succinct definitions of modelling and simulation are neatly provided from a systems engineering perspective in [169], quoted as follows:

What is modelling?

"Modelling is the process of producing a model; a model is a representation of the construction and working of some system of interest. A model is similar to but simpler than the system it represents. One purpose of a model is to enable the analyst to predict the effect of changes to the system. On the one hand, a model should be a close approximation to the real system and incorporate most of its salient features. On the other hand, it should not be so complex that it is impossible to understand and experiment with it. A good model is a judicious tradeoff between realism and simplicity."

What is a simulation?

"A simulation of a system is the operation of a model of the system. The model can be reconfigured and experimented with; usually, this is impossible, too expensive or impractical to do in the system it represents. The operation of the model can be studied, and hence, properties concerning the behaviour of the actual system or its subsystem can be inferred. In its broadest sense, simulation is a tool to evaluate the performance of a system, existing or proposed, under different configurations of interest and over long periods of real time."

This chapter so far has discussed modelling methodologies used for representing wireless CSMA/CA networks, classified broadly into analytical modelling and normative simulation approaches. This section now defines a middle ground 'conceptual simulation modelling' methodology, widely used in applied mathematics but not commonly found in the wireless modelling literature. This approach combines the beneficial characteristics of analytical and normative approaches. Each of the three research methodologies carries with it inherent assumptions about the network in question, with implications for the validity and usefulness of the findings that the methodology is able to generate.

This section is structured as follows. Sec. 3.3.1 explains conceptual simulation modelling. Sec. 3.3.2 explains the need for conceptual simulation models by discussing reasons for modelling and simulation. Sec. 3.3.3 then presents a direct comparison between analytic modelling, normative simulation and conceptual simulation modelling to clarify the novelty and to explain the unique features of the proposed conceptual simulation modelling methodology.

3.3.1 Conceptual Simulation Modelling

A conceptual model is a description, of the thing being modelled, describing the objectives, inputs, content and outputs by their relationships, together with the reasoning for assumptions and the simplifications that are made [215]. Essential requirements include: firstly, defining the system boundary (i.e., what aspects of the real system are to be included in the model and what is to be left out) and secondly determining the required detail to be included for each sub-section of the model and what features can be neglected. It should be asked: what impact will adding or neglecting details have on the results when performing simulation based on the conceptual model?

Conceptual model validity [225] is determining that firstly, the underlying theories and assumptions of the model are correct, and secondly, that the model representation of the problem entity and the model's structure, logic, and mathematical and causal relationships are realistic for the intended purpose of the model. A common method of validation involves simulating the model under well-understood conditions for comparison of the model outputs with the real system outputs [169]. This is obviously only possible in situations where known outputs exist. When using simulation as a design tool for new systems, this approach may not be possible.

All models are conceptual models, including the analytical (mathematical) and normative simulation models discussed. However, analytical and normative approaches represent opposite ends of the modelling spectrum. Analytical modelling, the traditional theoreticians' approach, assumes the system can be modelled as a series of solvable equations, representing ongoing processes. Normative simulation modelling, the traditional engineers' approach assumes that the interconnected parts of the system are too complicated to be represented by a series of equations and simulating their interaction will produce otherwise unforeseen emergent characteristics.

Conceptual simulation models present an approach somewhere in between analytical mathematical models and normative simulation. They have the potential to provide the flexibility and control of analytical models whilst enabling accuracy closer to normative simulation. Conceptual simulation models provide a structured approach to dealing with complexity with a mixed methodology that uses the strengths of problem structuring methods (PSMs) [179] to add value, extending their suitability to problem exploration and idea generation which traditional analytical or normative methods fall short of achieving. Fig. 3.12 represents the range of modelling methodologies considered.

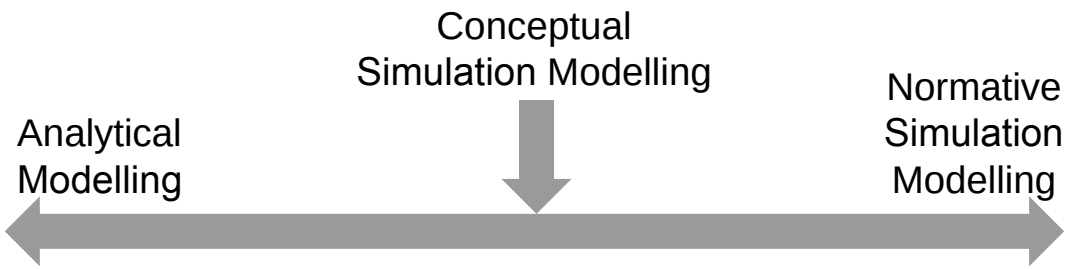


FIGURE 3.12. The range of research methodologies for wireless network.

3.3.2 Why Use Modelling and Simulation?

The reasons for building conceptual simulation models can be justified by identifying their purpose. The need for modelling and simulation stems from a desire to understand or replicate complex situations that may be impossible or improbable to observe and analyse due to scale, expense, accessibility and sensitivity [169]. Analysts rely on interpreting existing information. For some situations information may not exist, and modelling and simulation are tools that can be used to generate data. Simulation, modelling and analysis allow practitioners to gain an understanding of a system and its operations [169]. Often these approaches have the advantage that time can be compressed so that systems can be examined over long periods that are not practical for real-world investigations. Further, numerous tests can occur with the parameters of the model manipulated and changed to test different scenarios or hypotheses. A model / simulator can be built to generate the data required that may not otherwise be available, providing insight into unknown situations aiding design choices before implementation to ensure the best possible solution [215].

Modelling and simulation techniques are used in many familiar situations such as weather forecasting and financial prediction. Often their intention is forecasting or predicting future events, but there are many examples of simulation being used for other purposes. For example, in economics, simulation may be used to judge the desirability of proposed policy actions. Further, evacuations of buildings may be simulated for emergency planning, military simulations are used for both testing and rehearsal of various tactics, and business simulations may be used for the education of students. Epstein [80] identifies sixteen reasons to build models, other than prediction, quoted as follows:

1. "*Explain (very distinct from predict).*
2. *Guide data collection.*
3. *Illuminate core dynamics.*

4. *Suggest dynamical analogies.*
5. *Discover new questions.*
6. *Promote a scientific habit of mind.*
7. *Bound (bracket) outcomes to plausible ranges.*
8. *Illuminate core uncertainties.*
9. *Offer crisis options in near-real time.*
10. *Demonstrate tradeoffs / suggest efficiencies.*
11. *Challenge the robustness of prevailing theory through perturbations.*
12. *Expose prevailing wisdom as incompatible with available data.*
13. *Train practitioners.*
14. *Discipline the policy dialogue.*
15. *Educate the general public.*
16. *Reveal the apparently simple (complex) to be complex (simple).*"

Within the systems engineering context of developing complex technical systems, such as transport systems, supply chains, logistics networks, wireless communications, civil infrastructure and aerospace amongst others, some reasons for modelling and simulation are listed as follows.

- Provoke thoughts that lead to new ways of thinking (and to new models).
- Visualise the performance / operation of a system so others may understand.
- Explore design (ideas).
- Explore the effect of individual mechanisms.
- Formally explore the parameter space of proposed designs.
- Explore / find optimal parameters for proposed designs.
- Robustness / sensitivity testing of selected designs.
- Final normative test of proposed design, attempting to get as close as possible to a faithful representation of the world (expensive in terms of CPU and and man hours).

The above reasons for modelling and simulation illustrate different requirements of a model or simulation that depend on the application context. The level of complexity, build time, format and accuracy of the modelling or simulation approach chosen must reflect the requirement. The stated requirements, individually, could be achieved by a model positioned on the range depicted in Fig. 3.12. The final point in the list 'Final normative test' would, of course, fall at the normative simulation modelling end of the range, however, others such as 'Explore design' do not require full normative simulation. Some may be achievable by purely analytical methods but others require some sort of simulation (but not a full normative simulation), e.g., 'Visualise the performance'. Conceptual simulation models provide a structured approach to managing complexity and understanding a system.

Ideally, a simulator should model all relevant aspects of the system, be easy to modify, run fast, and produce accurate, reliable results. Determining the appropriate level of abstraction is paramount since the simulator's accuracy and speed are almost always inversely related [44, 253]. The value of a model or simulator should be judged on its ability to fulfil the original goal, i.e., to aid with decision making or understanding [153]. In a simulation environment, ideas can be tested risk-free, helping to build knowledge and understanding. Models can provide a powerful method of visualising and communicating ideas. Simulations can be used to provide a consensus testing different theories [215]. Achieving some or all of these goals could be an indication of the value of the simulation. To build a simulator of value requires time, expertise and expense [215]. The investment of research effort, time, money and resources in building a simulator in return for the value achieved has to be judged on a case by case basis. The appropriate level of detail must be determined by questioning if the model is sufficiently complex to address the relevant issues.

With all models and simulations, the limitations must be recognised. The data extracted can only be as good as the model. An understanding of the inherent assumptions is essential and the validity of the model for the particular purpose of use must be considered. An erroneous assumption, inaccurate input data, or errors in the model, will all inevitably lead to a loss of accuracy and usefulness.

3.3.3 Comparison of Methodologies

The following provides a comparison of the three modelling methodologies: analytical modelling, normative simulation modelling and the proposed conceptual simulation modelling (see the range illustrated in Fig. 3.12). In addition to the technical approaches, the inherent assumptions that a user makes when selecting their chosen research method (in relation to queuing networks) are compared (see Tab. 3.2). The purpose of this comparison is to highlight how a conceptual simulation modelling approach can provide a useful method for representing networks. Although the comparison (and this chapter) focusses on wireless networks, the ideas could be applied in other domains.

For clarification, when referring to analytical modelling (although there are many variations

of analytical queuing theory investigation and the term is used rather loosely), it supposes models represent the system variables by a series of equations designed to capture random events: the approach of Bianchi [29] would be an example. When referring to normative simulation modelling: it supposes a methodology that creates the most accurate representation of the system possible: NS3 [196] is an example. Conceptual simulation models present an approach somewhere in between analytical models and normative simulation (and hence between [29] and [196]).

Table 3.2: Comparison of Network Modelling Methods

Comparison Measure	*Analytical Modelling*	*Conceptual Simulation Modelling*	*Normative Simulation Modelling*
Granularity: Different levels exist [114]. The appropriateness of a macroscopic or a microscopic (i.e., fluid level or packet level) simulator must be determined. The decision will largely depend on the level of understanding required from the model and the data available when building.	Macroscopic model: considers packets flowing as a fluid. Fluid characteristics can be extracted such as average flow and density of traffic but the detail at the microscopic level is lost.	Reducing the complexity of the real system to develop a conceptual simulation model combines together some of the microscopic granular characteristics. The method however should still allow for characteristics to be extracted in relation to individual packets.	Microscopic model: the dynamics of single packets are delineated to enable consideration of characteristics of an individual packet. This provides a high fidelity description useful in detailed studies but is extremely computationally demanding.
Network topology representation.	Random distribution for a given network density.	Arbitrary location (randomly generated or specified) for chosen number of nodes, with interconnectivity and interference defined by prescribed rules.	Individual node positions defined.
Random event representation.	A probability distribution can be derived to define occurrences of random events.	Heuristics defining random events are built into the model.	Individual components of the model have their own characteristics and random elements. The overall behaviour of the model then depends on the interaction of the components throughout the simulation.

Table 3.2: Comparison of Network Modelling Methods

Comparison Measure	*Analytical Modelling*	*Conceptual Simulation Modelling*	*Normative Simulation Modelling*
Simulation / model Paradigm. See [171].	Process orientated paradigm: models the activity of the system as a series of processes e.g., arrival process, service process, routing process. The system boundary is defined by rates of movement of packets and numbers of nodes.	Activity orientated paradigm: generally the easiest to understand, as events are viewed chronologically, and therefore most appropriate for reducing the complexity. Simulation time might be discretised by breaking it down into increments negligible in relation to the length of the overall simulation. With each timestep, the simulator will look to update the state of each entity in the system.	Event orientated paradigm: eliminates timesteps where no events occur, instead, at the end of one event it moves directly to the next. An event list is used to store pending events, i.e., the arrival of new packets or the completion of service.
Ontology: To do with our assumptions about how the system being modelled is made up.	Assumes traffic through the system can be considered as a fluid level flow. Processes can be captured by a probabilistic model and output in terms of a probability distribution is sufficient.	Assumes the simplification that certain parameters of the system can be neglected. Arrival process can be captured by a probabilistic model.	The system is complex and needs to be represented as accurately as possible.

Table 3.2: Comparison of Network Modelling Methods

Comparison Measure	*Analytical Modelling*	*Conceptual Simulation Modelling*	*Normative Simulation Modelling*
Ontology: What it assumes to exist?	A function describing the distribution of inter-arrival times. A service times distribution function. A known number of servers. The capacity of the system, i.e., the maximum number of packets in the system including the one being serviced. A population whose size can be defined, i.e., number of sources of packets. A service discipline.	A knowledge of the parameters of the protocol that will be implemented. An approximate topology of the network that can be summarised as an adjacency matrix. A parameter space.	An accurate PHY and MAC layer model considered representative of real-world environments. A defined Network Topology. A traffic pattern with defined origin and destinations. A defined protocol. A sufficient amount of data to build a highly accurate model.
Epistemology: To do with our beliefs about how one might discover knowledge about the world.	Generate statistical data from which theory can be developed.	Generate data across parameter space. Generates a representation of phenomena occurring and data to support (not the same level of accuracy as normative modelling).	Accurate modelling will produce an accurate result.

Table 3.2: Comparison of Network Modelling Methods

Comparison Measure	*Analytical Modelling*	*Conceptual Simulation Modelling*	*Normative Simulation Modelling*
Epistemology: Is the research method concerned with finding out about objective facts and exact data relating to them? Or is the research method concerned with ideas and phenomena that have no external reality; i.e., phenomena that are interpretations by the researcher?	This methodology provides facts about the network. These, however, can be limited in their utility, and accuracy, for a dense network with complex topology. The results are likely to provide theoretical bounds on capacity, throughput, etc that may be significantly different from what is achievable in reality.	The methodology provides a method of investigating ideas and phenomena. The methodology aims to be an approach to generating accurate reliable data, however, it acknowledges that due to some simplifications there will be a degree of error.	This methodology is interested in precise facts. The method tests a clearly defined set of parameters and the data generated is accepted as accurate.

Table 3.2: Comparison of Network Modelling Methods

Comparison Measure	*Analytical Modelling*	*Conceptual Simulation Modelling*	*Normative Simulation Modelling*
Axiology: What is valued or considered good. This manifests itself in what the purpose or use of the models are, and who develops and uses the model (necessary information, source of information, users, purpose).	The basic high-level information such as arrival process, network size and service process are considered sufficiently represented by a defined probabilistic process. The analytical methodology can be used to produce statistical information concerned with overall network performance. The methodology can be used as a planning technique or an analytical technique for comparison with theoretical achievable performance and actual achieved performance.	A validation for the assumptions the methodology makes is made by strong supporting literature review. Using this methodology requires a demonstrated level of expertise in the application and support by data from literature from which the inherent assumptions in the model can be justified. The methodology can be used by researchers for the exploration of ideas.	The methodology takes the opinion that accurate models produce accurate information. It assumes that the component parts of the system can be accurately modelled from the available MAC and PHY layer data and that our understanding of the interrelation between component parts can be extrapolated for exploring new ideas for which no real data is available. The methodology is applied for thorough testing of known network designs.

Table 3.2: Comparison of Network Modelling Methods

Comparison Measure	*Analytical Modelling*	*Conceptual Simulation Modelling*	*Normative Simulation Modelling*
Axiology: Were the facts there before the researcher took an interest in them and are the phenomena being researched regarded as objects? Do values help determine what are recognised as facts and the interpretations drawn from them?	The methodology is objective and unbiased. The data generated provides insight from which further theories can be extracted.	The methodology is objective, however, a degree of bias may occur relating to the researchers' viewpoint. This can be minimised by providing strong literature grounding to support any assumptions. The results generated aim to identify phenomena and provide understanding of their origins.	The methodology is objective and unbiased. However where data is unavailable on certain details of models, assumptions will be made. This may introduce some bias. The data generated is assumed to be an accurate, comparable to data generated by a physical implementation.

Table 3.2: Comparison of network modelling methods.

Analytical modelling and normative simulation have many applications for which they are the most appropriate research methods of those discussed. It is a strength of these methods that they are unbiased and independent of the researcher. Applying these methods can lead to generalisable findings, however, the weaknesses of the methods is their focus on testing a specific setting and a lack of flexibility in generating new ideas (confirmation bias [194]). Further, there is no guarantee that knowledge generated is transferable to other examples [131]. It is a strength of normative simulation to be able to produce very accurate results about a specific topology, however, it is limited in its ability to test multiple different design options due to design complexity and simulation run time.

Developing a mixed methodology that combines aspects of analytical modelling, normative simulation and PSMs can use the strengths of each to overcome the limitations of the others. Mingers and Rosenhead [179] explain that what a PSM offers is a method of representing the situation that will enable users to bring clarity to the 'mess', identifying a potentially actionable common problem or issue within the system, and determining an amendment that will at least partially resolve it. They further explain that to do this a PSM must: "enable several alternative perspectives to be brought into conjunction with each other"; "be cognitively accessible to actors with a range of backgrounds and without specialist training, so that the developing representation can inform a participative process of problem structuring"; "operate iteratively, so that the problem representation adjusts to reflect the state and stage of discussion among the actors, as well as vice versa"; "permit partial or local improvements to be identified and committed to, rather than requiring a global solution, which would imply a merging of the various interests." The conceptual modelling methodology achieves these objectives. Through the process of developing a conceptual model, the perspectives of those involved are captured. The methodology does not attempt to explain the source of all uncertainty, recognising that attempting to model the situation exactly may not be possible and attempting to do so will result in a series of less controlled errors, which will be more difficult to understand in the simulation output. Researchers with different perspectives may form different opinions on the source of the uncertainty in the real system, however, collectively they define a model to capture that.

Rosenhead [216] identifies a benefit of PSMs to be a focus on managing, rather than reducing, complexity and thus aiding decision-makers in developing a comprehensive understanding of the situation and consensus on what the 'problem' is. The conceptual modelling methodology discussed can be argued as a method of developing an appreciation of the situation which may prompt further investigation. In line with the explanation in Ackermann [4], managing the complexity and thus being able to see the whole facilitates groups to achieve several benefits such as ensuring the situation is investigated from multiple perspectives and therefore enabling the development of a representative appreciation, broadening the number of alternatives generated, and finally allowing new options to emerge.

Applying a multi-methodological approach [131] to research in wireless networks, by develop-

ing a conceptual model combining analytical modelling and normative simulation features with PSMs, has several advantages. The model provides a description of complex phenomena, that otherwise may not be identified in the system. The model produced captures the users' perception and understanding and generates further understanding which is transferable to other examples. The method uses strengths of PSMs to overcome the limitations of analytical techniques. The results may contain an element of bias from the researchers' perspective, however, this can be managed and justified by identifying all assumptions made and supporting them with existing literature. A reduced-complexity conceptual model may fail to capture important detail in some complex situations: this will need to be assessed on a case by case basis but it should be recognised that for many complex systems, detailed normative simulation may provide data too complex to draw any useful conclusions from.

Due to the multidisciplinary nature of systems engineering projects (e.g., wireless systems), collaboration of researchers from a range of backgrounds with differing expertise, such as mathematicians, engineers and computer scientists (and in the case of wireless systems, MAC layer specialists, PHY layer specialists) is required. The reduced-complexity conceptual simulation model provides a simplification that is easy to understand. Further satisfying the definition of a PSM [179], as understanding develops, a conceptual model can similarly develop. The model can be adapted with the benefit of new knowledge and then re-tested. As a result, specific problem areas can be identified and solutions developed.

3.4 Conclusion

This chapter contributes to knowledge in the following ways:

CTK 3.1 Defining conceptual simulation modelling, for the purpose of investigating wireless networks and identifying how it can be beneficial as a research methodology.

CTK 3.2 Identifying the key assumptions and implications of each of the three defined methodological categories: analytical modelling, normative simulation and conceptual simulation modelling, for the purpose of better-informing researchers in their selection.

CTK 3.3 Identifying the unique features and capabilities each methodology offers in the development of complex systems.

Achieving the next-generation of wireless technologies will require more than just an incremental advance on their predecessors [10]. Rather, significant evolution is required in both the MAC and PHY layers. In this chapter three categories of modelling and simulation research method namely, analytical modelling, normative simulation and a proposed 'middle ground' conceptual modelling methodology, have been compared and their use in aiding the development of next-generation wireless networks discussed. It is acknowledged that a conceptual model is, as all models are, an approximation and therefore there will be a degree of uncertainty, however, if

the purpose of the model is to bring clarity to a complex, wicked problem [214], the model can still provide a significant contribution. As a result, the methodology can be considered as a PSM, i.e., a structured approach to dealing with complexity.

In light of these discussion the following Chapter 4 proposes a novel conceptual simulation modelling methodology for investigating performance characteristics of wireless networks. This chapter has provided the necessary background on CSMA/CA wireless networks and existing modelling and simulation techniques for understanding the contributions in the following chapters.

CHAPTER 4

A Conceptual Simulation Modelling Methodology for the 802.11 CSMA/CA MAC Protocol

This chapter presents a conceptual simulation modelling methodology for representing wireless CSMA/CA networks. The methodology is designed to aid researchers in the environment of a corporate research and development lab exploring next-generation wireless technologies, with the goal of developing innovations and IP that can bring commercial success. The methodology is shown to produce results comparable to those in existing literature and a high degree of correlation with results produced via normative simulation (i.e., NS3) for some special test cases. Here, the conceptual simulation modelling methodology is presented in its most simple format. Later chapters of this thesis extend the model to fit various detailed scenarios of interest.

The methodology is designed for simulations of network traffic for a large number of demand scenarios. This is computationally demanding, so to reduce the CPU time needed, the simulation methodology implements a simplification of the MAC protocol. The most significant simplification is in the pre- and post- stages of packet transmission (RTS, CTS, ACK). These are not explicitly modelled, rather, their influence is captured by the introduction of a *Countdown* state to the protocol. This is heuristically designed to mimic the relevant complexity that is omitted by simplifying other processes. A discrete timestep is used as opposed to a discrete event methodology for simplicity. The model captures the state of links (i.e., possible transmissions) for arbitrary topologies and cycles through the model's protocol states, pending an integer number of timesteps in each, while respecting the interference constraints imposed by the wireless medium.

The chapter is structured as follows. Sec. 4.1, echoing the PSM thinking of Sec. 3.3, provides an overview of the requirements discussing purpose, intended users, how the methodology is suited for a successful study and how validation of such a methodology is achieved. Sec. 4.2

presents the conceptual simulation model for representing wireless CSMA/CA networks. Sec. 4.3 presents an ensemble simulation methodology for implementing the conceptual model. Sec. 4.4 provides a number of example simulations comparing the conceptual simulation model with NS3 for validation purposes. Sec. 4.5 gives conclusions.

4.1 Requirements

This section is structured as follows. Sec. 4.1.1 explores the process of conducting a successful modelling and simulation study and explains how the proposed conceptual simulation modelling methodology meets these requirements. Sec. 4.1.2 discusses how the methodology can be considered valid, and why it is useful.

4.1.1 Conducting a Successful Modelling / Simulation Study

The methodology presented in this chapter is designed to aid researchers investigating next-generation wireless technologies in the context of a corporate research and development lab. Although the conceptual simulation modelling methodology presented produces quantified numeric output, it is not designed for normative testing like other simulators common in the wireless industry (e.g., NS3 [196] or OPNET [199]). Instead, it provides a tool for developing a more fundamental understanding of CSMA/CA networks and newly proposed technologies for these networks, in real-world environments, bringing together expertise from different backgrounds i.e., wireless experts, engineers, mathematicians. The apparently simple methodology reveals the complexity of the interacting component parts of CSMA/CA networks (reasons for modelling were discussed in Sec. 3.3.2 p.81).

The model is designed parsimoniously [253], capturing the replicable detail of the CSMA/CA communications process and neglecting idiosyncratic, non-replicable parts, simplifying these into an error parameter. The parsimonious approach distinguishes important interactions of the CSMA/CA process from other negligible background details, producing a model that enables understanding.

Law (2009) [153] explains a seven-step approach for conducting a successful modelling / simulation study. The steps are presented and how they are achieved in this thesis expanded upon in relation to the conceptual simulation methodology presented.

1. “Formulate the problem.”

   The problem, briefly, is that CSMA/CA networks are complicated. There is a need to understand the potential performance impact of proposed next-generation wireless technologies which is difficult to achieve with existing methods. Analytical models fail to capture the required level of detail. Normative representation of CSMA/CA networks is challenging, time-consuming and computationally demanding, and further, the detail of some proposed

next-generation technologies is not yet clear making normative representation impossible. Thus, new methodologies are required.

2. “Collect information / data and document assumptions.”

   Background information on the structure of the system being modelled and its operating procedures has been provided and a literature review of relevant approaches to modelling and simulation was conducted (see Chapter 3). Throughout this chapter, the parameters of the 802.11 CSMA/CA system and the model are presented and explained. Further, inherent assumptions and uncertainty in the model are documented and discussed. Quantitative validation of the new model is conducted by comparison of simulation results with more normative simulators which are already widely used in the industry and have undergone their own validation (i.e., NS3 [196]). This thesis provides the necessary detail for an individual to be able to build their own version of the model and recreate the application-specific studies presented in the chapters that follow.

3. “Are the documented assumptions of the model valid?”

   Extensive documentation of all concepts and assumptions along with justification from supporting literature and further data generated give credibility to the model. A detailed description of how the conceptual simulation methodology presented is assembled and how the component parts of the model interact is provided.

4. “Program the model.”

   Details of the component parts of the model and their relationships are given. Further, meta code is provided in the appendix of this chapter (App. 4.A.1 p.125). The simulation model is constructed in the activity orientated paradigm [171] with time broken into tiny increments. For this thesis, the conceptual simulation model was programmed in, and all the results presented produced in, Matlab [170]. However, it could be implemented in any modern programming language.

5. “Is the programmed model valid?”

   The most definitive test of a simulation model’s validity is establishing that its output data closely resemble the output data that would be observed from the actual system. Results overlaying the presented conceptual simulation model with results generated via normative simulation in NS3 [196] are provided. NS3 is widely used in the wireless industry and considered a valid representation of CSMA/CA systems. The presented results show the conceptual simulation model’s ability to closely replicate results generated in NS3. Further, the commonality between results of presented studies and existing literature gives confidence in the methodology as a tool for representing the 802.11 CSMA/CA system.

6. “Design, conduct, and analyse experiments.”

Chapters 5, 6 and 7 demonstrate investigations evaluating the performance impact of proposed next generation technologies introduced into wireless CSMA/CA networks.

7. "Document and present the results."

   Presenting results visually is key in providing understanding. Critique of results and discussion in relation to existing literature is provided for every experiment. Results in line with existing literature give confidence in the validity of the conceptual simulation modelling methodology and in the subsequent novel findings and contributions to knowledge generated.

This seven-step approach from Law [153] provides an overview of the process of developing the conceptual simulation modelling methodology this chapter goes on to present. In particular, it reinforces that a reduced-complexity abstract model can be a useful tool providing assumptions are documented and justified.

4.1.2 Verification, Validation and Usefulness

Verification is the process of ensuring that the proposed model design has been correctly transformed into a computer model [69]. Validation is the process of ensuring that the model is sufficiently accurate for the purpose at hand [44]. Thus, verification has a narrow definition relating to the accurate construction of the model whereas validation is much broader. The usefulness of a model is judged on its ability to fulfil a purpose.

Validation is the process of determining whether, for the specific objectives of the investigation, the model accurately represents the system [153]. During the process of developing a model, there are certain validation activities that are to be conducted. These include [150] conceptual validation (are assumptions appropriate?), logical validation and verification (are the components of the model and their relationships a suitable representation of the real system?), experimental validation (do results generated by the model reflect real system data?), operational validation (do experts accept the findings of the model and do they use them to generate new knowledge?), and data validation (is the accuracy and appropriateness of the data acceptable for the cost and effort of building the model?). The following summarises some general perspectives on validation [153]:

- A 'valid' model can aid informed decisions, producing similar results, but in a more feasible and cost-effective way, than experimentation with the system itself.
- In some situations, it is possible to collect data on an existing system that can be used for validating a model. This is, of course, dependent on a current version of the system existing. The complexity of the validation process is dependent on the complexity of the modelled system.

- A model can only ever be an approximation of the complex system it is built to represent, regardless of time, money and effort invested in representing the system normatively. Hence, absolute model validity can never be achieved. Model validity is not consequently related to the development cost and further attempting to increase the validity of the model beyond a certain level can involve significant effort and expense that is not necessarily rewarded with significantly better insights or decisions. Box famously stated [35, 36], "All models are wrong, but some models are useful". He further posed the real question of interest, "Is the model illuminating and useful?".

- Validation is part of the model development process, not just something to attempt after the model has been completed.

The remaining sections of this chapter introduce the conceptual simulation modelling methodology and demonstrate that it can provide a valid representation of the 802.11 CSMA/CA system. A detailed description of the construction of the model, its parameters and simulation methodology is given. The chapters that follow demonstrate the usefulness of the model by applying it to problem situations.

4.2 Conceptual Simulation Model

This section presents two versions of a conceptual simulation model which capture the behaviour of wireless CSMA/CA networks and can be used to study proposed next-generation technologies. The models progress through discrete time and capture the state of links (i.e., possible transmissions) connecting each pair of nodes in a network. The link interactions are defined by a series of matrices (see Sec. 4.2.2). A model suitable for modelling CSMA/CA considering ideal conditions (i.e., perfect PHY, zero propagation loss and instantaneous carrier sensing) is presented consisting of five states. The model is then further refined for modelling CSMA/CA considering imperfect conditions, simplifying it to four states.

The cycle begins with the *Dormant* state. If a link has a packet queued to send, it senses the channel. If clear (i.e., if it were to transmit along the link it wishes to transmit along, to the best of the node's knowledge this would not cause a collision) it proceeds to the *Countdown* state, remaining there for a period randomly selected from a uniform distribution. After this delay, the link continues the cycle and moves to *Transmit* for the duration of the packet.

If the transmission completes collision-free the queue is decremented, a successful transmission is recorded and the link moves to the *Listen* state (five-state model only) where it waits for a fixed period before it returns to the *Dormant* state. If at some point during the transmission, another transmission is attempted along a link unable to transmit simultaneously due to interference, a collision is recorded. Collisions during the transmission are only recognised at the end of the *Transmit* state. In this instance the collided links enter a *Backoff* state for a

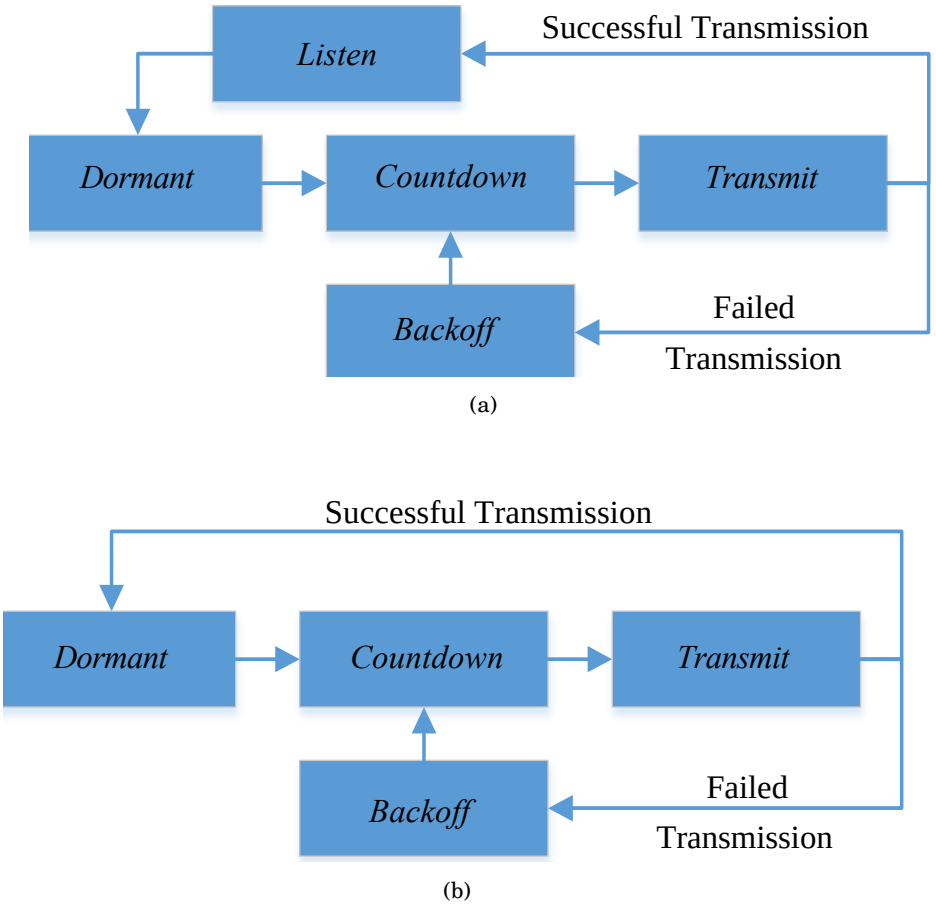


FIGURE 4.1. State diagram for each link of the conceptual simulation model. (a) Five-state and (b) four-state models. Dormant: node senses channel. Countdown: random time delay between sensing channel is free and beginning transmission. Transmit: a packet is transmitted. Backoff: random time delay following collisions. Listen: (five-state model only) time delay following a successful transmission.

random duration, again selected from a uniform distribution, before reassessing the channel in an attempt to resolve the conflict.

When the *Backoff* stage completes, the link listens to the channel, and if clear, enters *Countdown* and proceeds through the cycle again as above; if the channel is busy, then the *Backoff* state repeats with a new random duration.

The conceptual simulation model presented in Fig. 4.1(a) is suitable for modelling wireless CSMA/CA systems assuming ideal conditions (i.e., perfect PHY, zero propagation loss and instantaneous carrier sensing). Fig. 4.1(b) presents a further simplification of the model neglecting the *Listen* parameter, such that when a packet is successfully transmitted the link state immediately returns to *Dormant*, reducing the model from five states to four. Validation of the two models is given throughout Sec. 4.3.

The four-state conceptual simulation model is not suitable for modelling the behaviour of the 802.11 MAC protocol when considering ideal conditions. However this thesis argues that considering ideal conditions is, of course, unrealistic and therefore uninteresting. Fig. 4.2 shows a simple example network and illustrates how the four-state model cycles through states over time.

4.2.1 Conceptual Simulation Model Parameters

For both the five-state and four-state models, *Backoff* and *Countdown* counters are defined by a uniform distribution from which they are randomly generated on every occasion the model enters either of the two states.

The *Dormant* state has no time parameter associated with it. The model remains in the *Dormant* state for two reasons: either the link has no packet to transmit, or the channel is sensed busy and the link must wait before transmitting. The time spent in *Dormant* is therefore not dictated by the parameters of the model, but rather by the traffic arrival process and the channel environment.

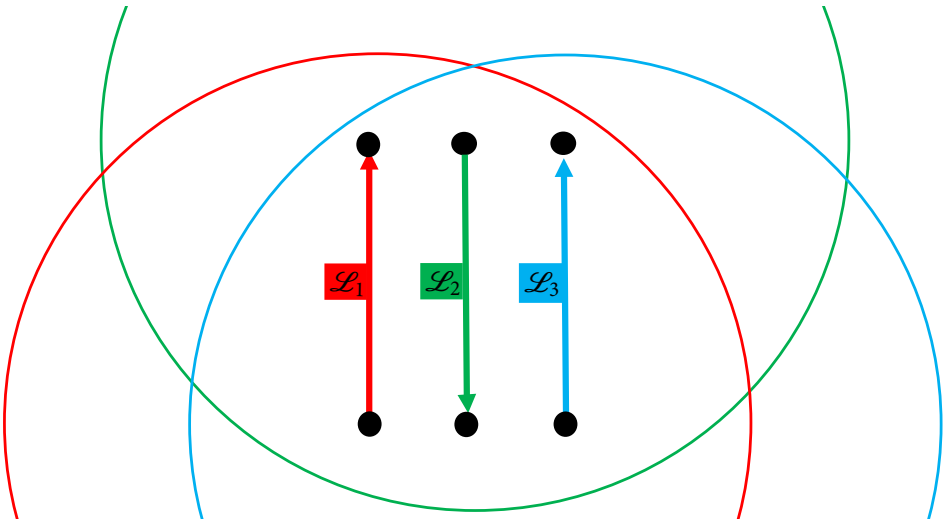
The *Transmit* state represents the transmission of a single packet, and its duration represents the time taken to transmit a single packet. For all simulations that follow, a fixed packet length is assumed and hence a fixed duration for the *Transmit* state.

When a successful transmission is complete the link waits in the *Listen* state for a fixed period of time (five-state model only). The *Listen* state captures the delay between consecutive packets transmitting on a single link that is characteristic of the 802.11 protocol. The five-state model will perform exactly the same as the four-state model by setting the *Listen* parameter equal to zero.

The *Backoff* state is designed to resolve collisions. When a collision occurs, links enter the *Backoff* state and wait for a randomly selected time before attempting to access the medium again. For simplicity, this time period is always chosen from the same uniform distribution and the timer decrements with every timestep of the simulation. Only at the end of the *Backoff* state is the channel sensed again. If then the channel is busy the link remains in *Backoff* selecting a new time from the same uniform distribution. For all simulations that follow, a fixed uniform distribution of the *Backoff* state is used.

Countdown is an artificial parameter used to introduced uncertainty in the carrier sensing process of the model. The upper and lower bounds of the parameter should be set dependent on the chosen level of carrier sensing error. Assuming ideal conditions, $Countdown_{\min}$ and $Countdown_{\max}$ should be set to zero. When set to zero, the carrier sensing procedure of the model will be perfect, such that links within carrier sensing range will always be aware of each other's transmissions and collisions will never occur.

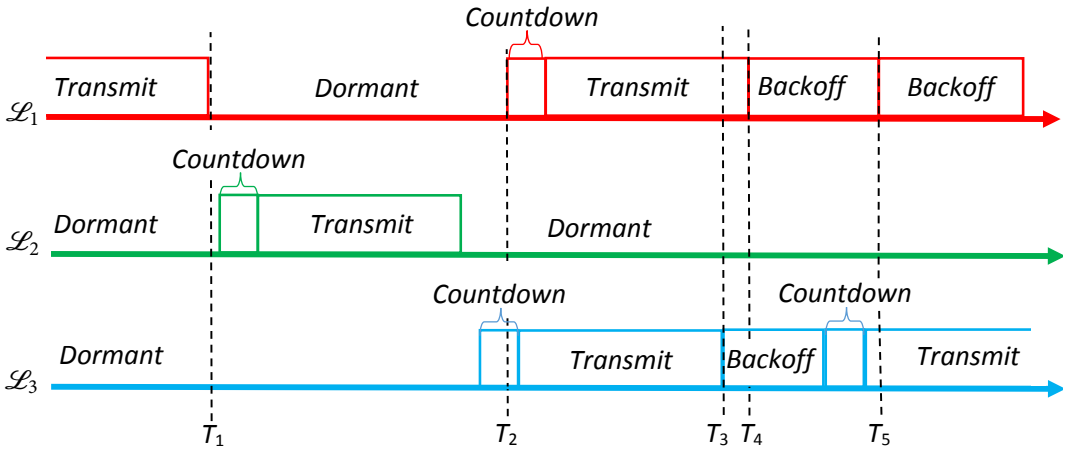
An explanation of how each of the parameters is calibrated in the simplified model is given in Sec. 4.4, along with validation, demonstrating the conceptual simulation model's ability to



(a)

$$\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

(b)



(c)

FIGURE 4.2. Link-centric conceptual simulation model. (a) A six-node three-link topology: circles indicate interference around the transmitter of the corresponding colour link. (b) A link collision matrix $\mathbf{E}$. As all links interfere with each other the 3×3 matrix is full of 1's except on the diagonal. Only one link may *Transmit* at any one time. (c) Illustrates the four-state model cycling through states as time progresses to the right. Key times are indicated. At T_1, $\mathscr{L}_1$ finishes *Transmit*. Almost immediately after, a packet arrives in $\mathscr{L}_2$'s queue; it senses the channel clear and so enters *Countdown* then *Transmit*. At T_2, $\mathscr{L}_3$, having previously sensed the channel clear is in *Countdown*. A packet arrives in $\mathscr{L}_1$'s queue; it too now senses the channel clear and also enters *Countdown*. Both then proceed to *Transmit*. At T_3 and T_4, $\mathscr{L}_3$ and $\mathscr{L}_1$ respectively realise their transmissions have collided and so enter *Backoff*. At T_5, $\mathscr{L}_1$ must *Backoff* again as $\mathscr{L}_3$ is in *Transmit*.

replicate wireless network characteristics and identifying the appropriate *Countdown* parameters for different levels of carrier sensing error.

It is important to note that the model parameters are not independent. The behaviour of the model largely depends on the ratio of lengths of different parameters. When halving the internal timestep of the simulator or doubling the length of every parameter, the characteristics of the simulator remain broadly the same.

4.2.2 Network Representation: Link-Centric Models

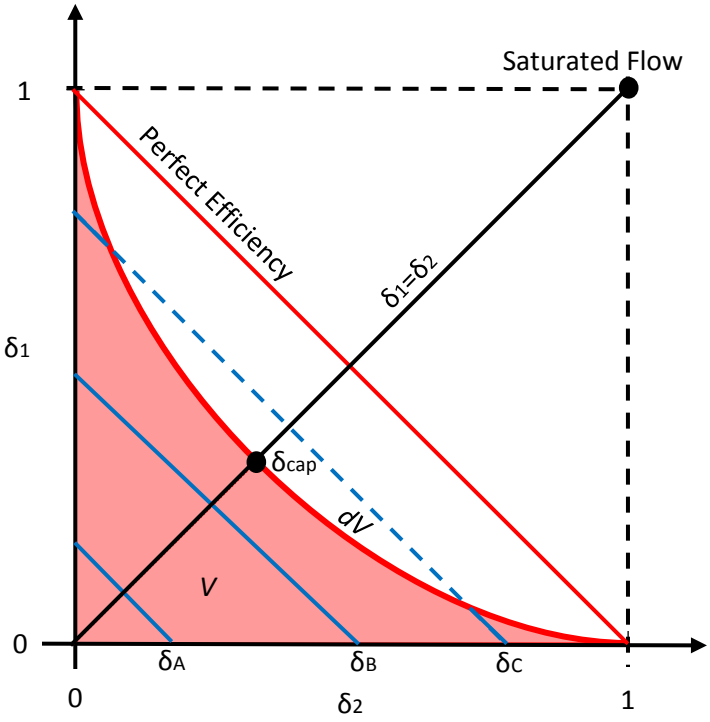
The approach of the conceptual simulation methodology is to capture the network via a link-centric model (see Fig. 3.8 p.71). The network topology is represented by a set of unidirectional links indicating the possible single-hop flow of data between nodes. Fig. 4.2(a) shows a three-link example. A binary collision model is assumed. A collision matrix $\mathbf{E}$ is used to describe interactions between the n links in a network, with entries E_{ij}, $1 \leq i,j \leq n$. If $E_{ij} = 1$, then link i cannot transmit simultaneously with j, whereas if $E_{ij} = 0$ then link i can transmit simultaneously with j. Fig. 4.2(b) shows the simple 3×3 collision matrix $\mathbf{E}$ for the three link example. Interference (and hence the collision matrix) may be, but are not necessarily, symmetric: that is, if link i collides with link j, then link j may or may not collide with link i, i.e., both $E_{ij} = E_{ji}$ and $E_{ij} \neq E_{ji}$ are possible.

Further to defining the topology and collision matrix $\mathbf{E}$, Fig 4.2(c) illustrates how using the four-state conceptual simulation model, each of the three links in the network cycle through the various states as time progresses.

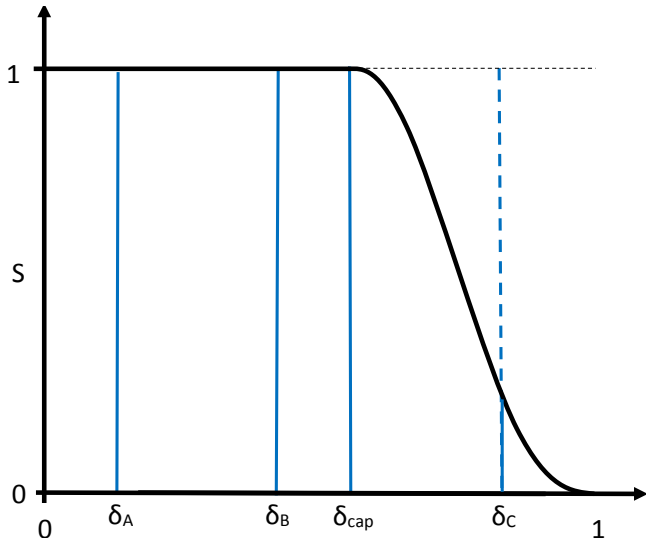
4.3 An Ensemble Simulation Methodology

In contrast to most literature which focuses on saturated demands, this methodology seeks to determine the capacity region. The aim is to identify combinations of links' demand intensity vectors $\boldsymbol{\delta}$ that can be met by the system. Demand intensity vectors are non-dimensional and range from 0 to 1 where 1 is the maximum transmission rate (or goodput i.e., the application-level throughput or number of useful bits delivered per unit time excluding failed transmissions and packet overheads) of a single link transmitting alone in a clear channel (i.e., the point a single link saturates the channel).

Fig. 4.3(a) illustrates an example capacity region for $n = 2$. However the capacity region can be identified for any number n of links. Numerical measures may be extracted to summarise the capacity region characteristics. Firstly, consider the line of equal demand $\delta_1 = \delta_2 = ,..., = \delta_n$. The value δ_{cap} identifies the maximum value (of $\delta_1 = \delta_2 = ,..., = \delta_n$) which is within the capacity region. Secondly, the proportion of the demand space $[0,1]^n$ that is within the capacity region can be measured and denoted V. Additionally, defining total demand intensity $\delta := \delta_1 + \delta_2 + ... + \delta_n$ then $\delta = \text{const.}$ describes sections (simplexes) through the demand space where total demand is



(a)



(b)

FIGURE 4.3. (a) Capacity region performance metrics for two links. Each demand intensity δ is varied from 0 to 1, i.e., the maximum transmission rate of a single link transmitting alone in a clear channel. The shaded area indicates the capacity region (of volume V) and the curved red line dV is its boundary. The concave capacity region indicates a loss of efficiency (and goodput) when links share the channel. The point δ_{cap} is the maximum value which is within the capacity region of the line $\delta_1 = \delta_2$. Blue lines perpendicular to $\delta_1 = \delta_2$ indicate sections of equal total demand $\delta_{\text{A}}, \delta_{\text{B}}, \delta_{\text{C}}$. The proportion of these sections within the capacity region is denoted S. (b) Plots S against δ.

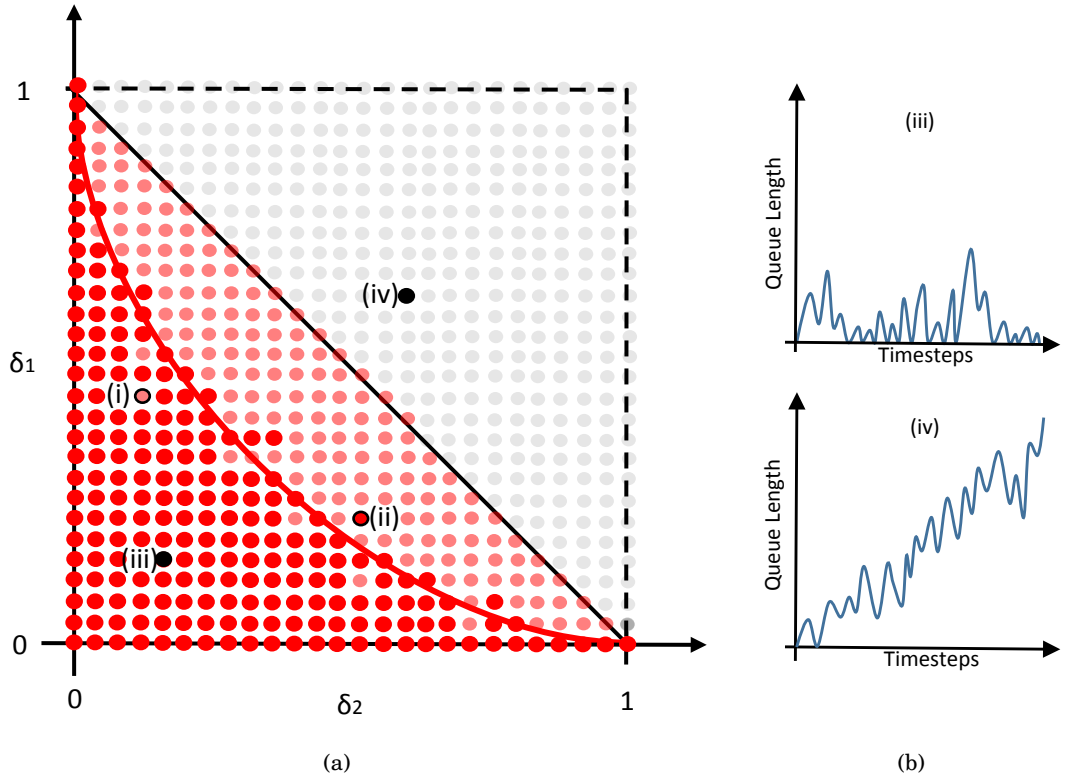


FIGURE 4.4. (a) Illustrates discrete simulation points. Solid red markers are assessed inside the capacity region. Faded red markers are assessed outside the capacity region, i.e., one or more queues is growing without bound. Faded grey markers are outside the maximum efficiency capacity region, i.e., both demands could not be met even considering a perfectly efficient system. A few individual points which have been misclassified are visible on either side of the capacity region boundary, however, across the full ensemble, the capacity region boundary is relatively stable. (i) and (ii) highlight two examples of points which have been misclassified. Points (iii) and (iv) identify simulations inside and outside the capacity region respectively. Part (b) illustrates how the queue length changes as the simulation progress for one link's queue. Simulations are judged to be outside the capacity region if the mean gradient of any links queue is greater than the threshold value 0.0001.

equal. It is therefore possible to count (as a function of δ) the proportion of the corresponding area ($n-1$ volume) that is within the capacity region, and denote this quantity S. It is hence possible to plot S against δ to examine the profile of the capacity region with increasing total demand (Fig. 4.3(b)).

To identify the capacity region, the methodology is to perform large numbers of simulations of different demands and test individually if each is inside or outside the capacity region. For a fixed set of demands, this judgement is made by observing the development of each link's queue length (i.e., number of packets queued) for each timestep of the simulation. If the demand is

being met the queue will remain stable and the simulation is inside the capacity region. If the queue is growing, the demand is not being met and the simulation is outside the capacity region.

Fig. 4.4(a) illustrates how by simulating ensembles of different demand vectors, assessing if each individually is meeting demand, the capacity region can be identified. The method is described for n links as follows. Consider a network of n links each with a demand vector $(d_1, d_2, ..., d_n)$. Let d_{sat} be the maximum transmission rate of a single link transmitting alone in a clear channel. Each of the demands $d_1, d_2, ..., d_n$ can be varied independently from 0 to d_{sat} in a defined number of increments K. It can therefore be considered that $\delta_1 := d_1/d_{\text{sat}}$, $\delta_2 := d_2/d_{\text{sat}}$,...,$\delta_n := d_n/d_{\text{sat}}$ denote non-dimensional demand intensities that range from 0 to 1. All possible demand intensity combinations can be exhaustively simulated resulting in $(K+1)^n$ individual simulations. The duration of each simulation is set at D/d_{sat} i.e., the time needed for one transmitter transmitting alone in a clear channel to transmit D data packets without collisions (typically $D = 10,000$ or greater is used throughout this thesis).

In the example, Fig. 4.4(a), $n = 2$ and $K = 25$ resulting in $26^2 = 676$ simulations. Of course, for this symmetric example set-up, half the points simulated are combinations of demand greater than the efficiency maximum (Fig.4.3(a)). Increasing the number of increments K will increase the resolution of the resulting capacity region (typically $K = 40$ is used throughout this thesis).

Fig. 4.4(b) shows examples of the queue length against time. Simulations have finite run time and the queue length is dependent on traffic arrivals and service rates. It is possible that even for a demand that can easily be served, in the final timesteps of the simulation, a number of packets arrive and not be served before the simulation finishes. This makes the gradient of the queue over time appear greater than zero (i.e., the queue growing). To avoid this problem, long simulation run times are desirable and a small tolerance is built into the gradient threshold. Simulations are judged to be outside the capacity region if the mean gradient of the observed queue length over time (Fig. 4.4(b)) is greater than a threshold value of 0.0001 for any single link's queue. This heuristic value was determined by trial and error.

Fig. 4.5 presents an example of capacity region simulation results of two demands in competition that illustrate why the growth threshold value is needed. All parameters of the simulations are constant except for the threshold value. In this plot $n = 2$ and $K = 40$ hence $1,681$ demand combinations are simulated. The threshold value has a significant impact on the number of points judged inside the capacity region.

This method of determining if a simulation is within the capacity region is, of course, a heuristic approach. The method was further 'sense checked' by comparison with other approaches. For example, consider a situation where multiple transmitters are sharing the channel. It is possible to observe the inter-service time of transmitted packets by a single transmitter over the period of a simulation. Mean inter-service time is equal to $1/\mu$ where μ is the service rate. If the arrival rate is greater than μ the demand is not being met. By performing a series of simulations that incrementally increasing δ, it is possible to identify the value of δ where μ can no longer

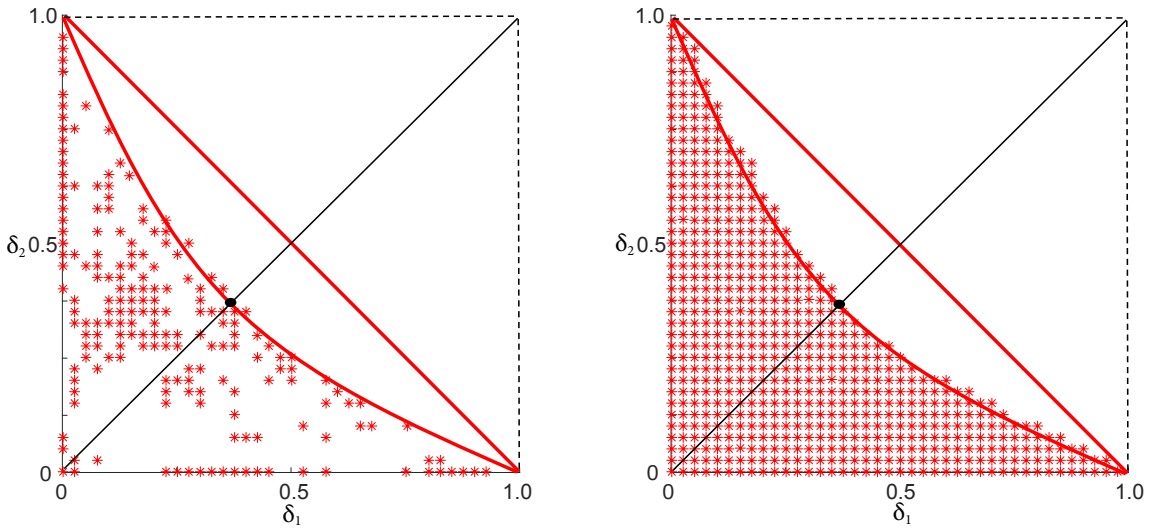


FIGURE 4.5. Simulation results demonstrating the need for a tolerance in the threshold used for judging if demand combinations fall inside or outside the capacity region. The threshold for judging if the queue of a simulation is growing over time is adjusted for the two simulations shown, with all other parameters kept the same. (a) Simulations are judged to be outside the capacity region if the mean gradient of the observed queue length over time is greater than zero. With this threshold (i.e., without any tolerance), only when the simulation ends with empty queues are they judged inside the capacity region. Due to Poisson traffic arrivals, in the final timesteps packets could arrive in the queue and not be served, even if the system is easily capable of serving the given demand rate. (b) Simulations are judged to be outside the capacity region if the mean gradient of the observed queue length over time is greater than 0.0001. Introducing this tolerance does not change the boundary of the capacity region, however, now demand combinations that are being met by the system are much more consistently judged inside the capacity region.

equal arrival rate. Increasing δ beyond this, μ remains constant (at its maximum). The value of δ for which μ reaches its maximum indicates the capacity region boundary. Using this method of assessing if a demand is within the capacity region and that of assessing queue evolution over time identify the same capacity region boundary. Although there is potential for some individual demand combination to be misclassified, overall the classification is found to be robust across the ensemble of discrete demand combinations described.

4.4 Parametrisation

To demonstrate the conceptual simulation models validity, a series of experiments are compared with NS3. The section that follows considers networks of n links transmitting within carrier sensing range of each other, i.e., all transmitters and receivers can hear each other. Each example,

therefore, uses an $n \times n$ collision matrix $\mathbf{E}$ full of ones except along the diagonal, which is zeros (a link cannot collide with itself), to describe interactions between links.

The remainder of the section is structured as follows. Sec. 4.4.1 explains how the level of carrier sensing error (ε) in the CSMA/CA protocol is controlled in NS3. Sec. 4.4.2 explains how the fixed number of timesteps assigned to the *Transmit* and *Listen* parameters were chosen. Secs. 4.4.3 explains how the parameters of the uniform distribution defining the *Backoff* and *Countdown* state are chosen. Secs. 4.4.4-4.4.6 provide a series of validation experiments by comparison with NS3.

4.4.1 Controlling Error Level In NS3

Consider simulating two links at saturated demand in competition for channel access via the CSMA/CA protocol using NS3. Under ideal conditions, i.e., perfect PHY layer, zero propagation loss and instantaneous carrier sensing, no collisions will occur at any stage of the protocol. This represents ideal CSMA/CA. For the validation experiments that follow, three levels of carrier sensing error (i.e., the probability that an RTS generated will not result in a complete CSMA/CA handshake with CTS, data packet and ACK) are examined; 0, 0.03 and 0.06 (denoted as ε_{min}, ε_{mid} and ε_{max}). These values relate to the bounds of the range derived by Bianchi [29] (see Sec. 3.2.4 p.72) considering the probability that in the CSMA/CA protocol, the backoff counter of two competing links will decrement to zero at the exact same moment based on the initial and maximum contention window, hence leading to a collision.

For this validation, to control ε in NS3, the propagation loss between each pair of links is manipulated. By observing the log of events in the NS3 simulator, it is possible to identify the impact of introducing propagation loss on carrier sensing error. Consider two links in NS3 with a set propagation loss to each other's transmitters and receivers (set using the NS3 matrix propagation loss model). Simulating saturated traffic for a period of time long enough to transmit $100,000$ packets one can identify ε, that is the number of RTS transmissions generated which do not lead to a successful CTS, data packet and ACK as a proportion of the total number of RTSs generated. By performing an ensemble of simulations, continually adjusting the propagation loss, the propagation loss required to achieve a mean carrier sensing error of $\varepsilon = 0.025, 0.030, ..., 0.060$ was determined. Fig. 4.6 plots propagation loss for $\varepsilon = 0.025, 0.030, ..., 0.060$.

For each experiment, the level of ε in NS3 is set between all links in the network to ε_{min}, ε_{mid} and ε_{max} by defining the propagation loss between each pair of links as discussed above. The level of ε can be replicated in the simplified conceptual model by controlling the *Countdown* parameter. Before discussing controlling ε however, first, the method by which the fixed parameters of *Transmit* and *Listen* are set is discussed.

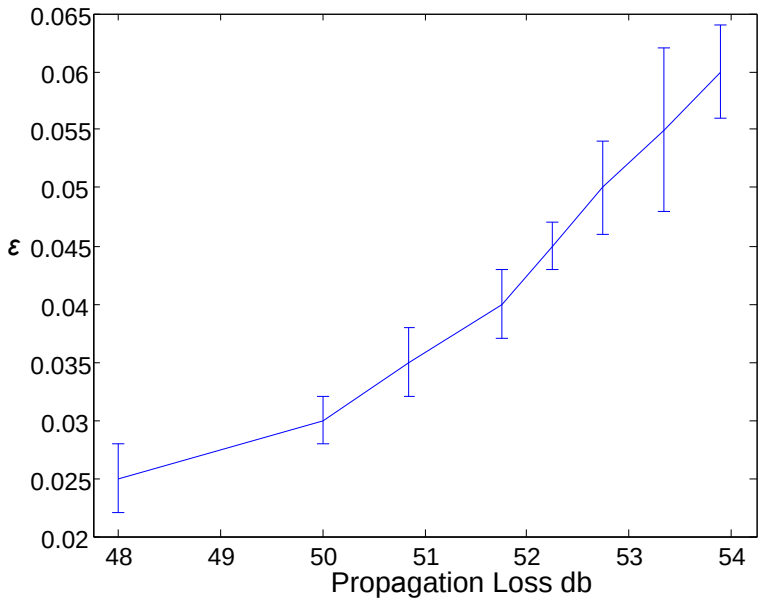


FIGURE 4.6. Carrier sensing error (ε) probability against propagation loss in NS3. The mean and range of 10 repetitions at $\varepsilon = 0.025, 0.030, ..., 0.060$.

Table 4.1: NS3 802.11b parameters [197].

Slot Time	$20\mu s$
MAC Header	272 bits
PHY Header	128 bits
ACK	112 bits + PHY header
RTS	160 bits + PHY header
CTS	112 bits + PHY header
SIFS	$10\mu s$
DIFS	$50\mu s$
CTS Timeout	$300\mu s$
ACK Timeout	$300\mu s$
W_{min}	31 Slots
W_{max}	1023 Slots
Retry Limit	7

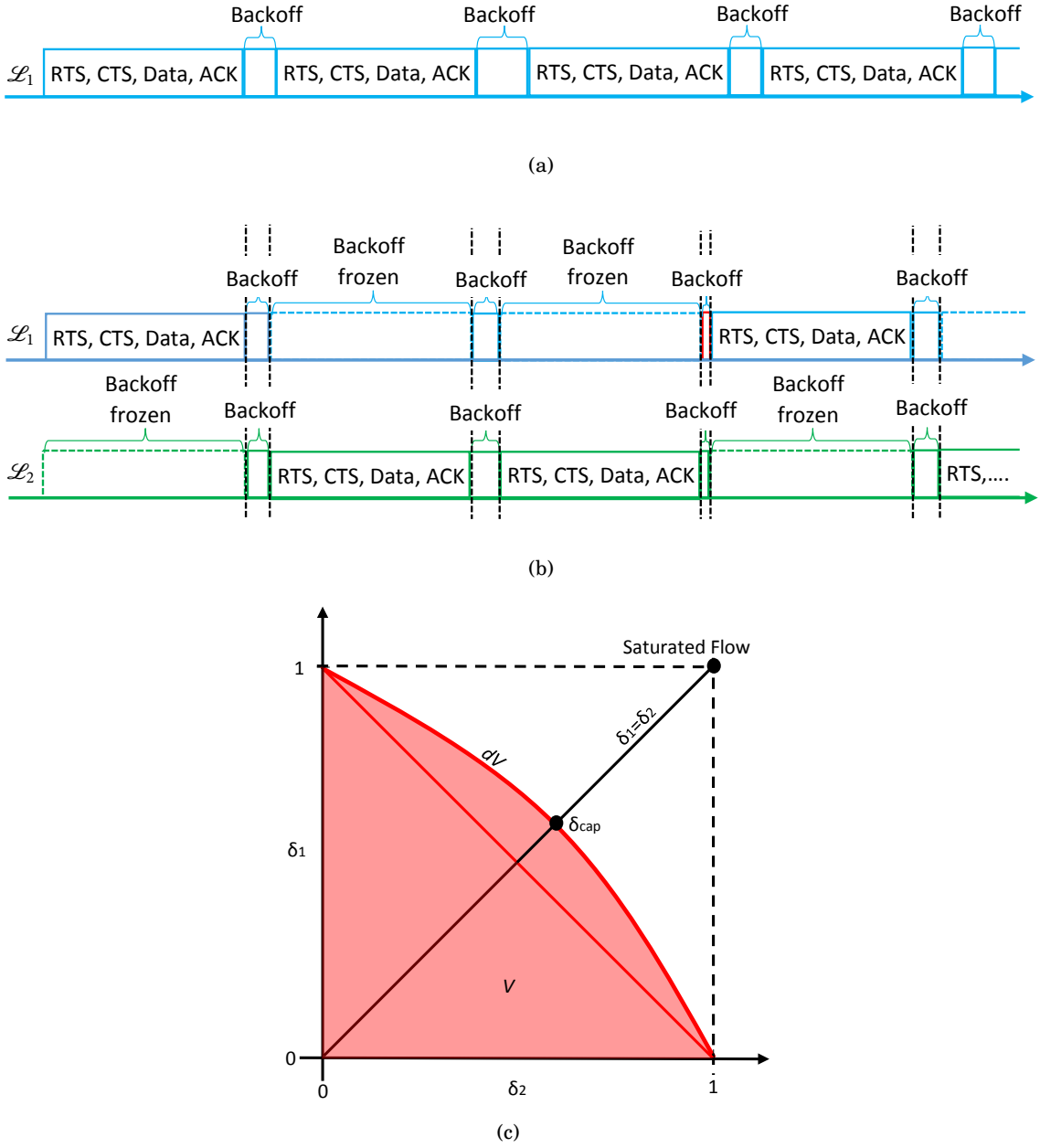


FIGURE 4.7. An illustration of why CSMA/CA can lead to a convex capacity region. (a) Shows a single saturated demand transmitting alone. Time proceeds to the right. (b) Shows two saturated demands sharing the channel without collisions. After successfully completing a transmission including RTS, CTS, data packet and ACK, a CSMA/CA transmitter must backoff for a period of time chosen from a uniform distribution, even when transmitting alone in a clear channel. If two transmitters are sharing the channel, the transmitter whose backoff counter decrements first will begin transmitting, hence the average delay between any two packets transmitting is reduced. The second transmitter's backoff is frozen for the duration of the transmission and resumes decrementing once it senses the channel clear again. (c) Illustrates the resulting convex capacity region.

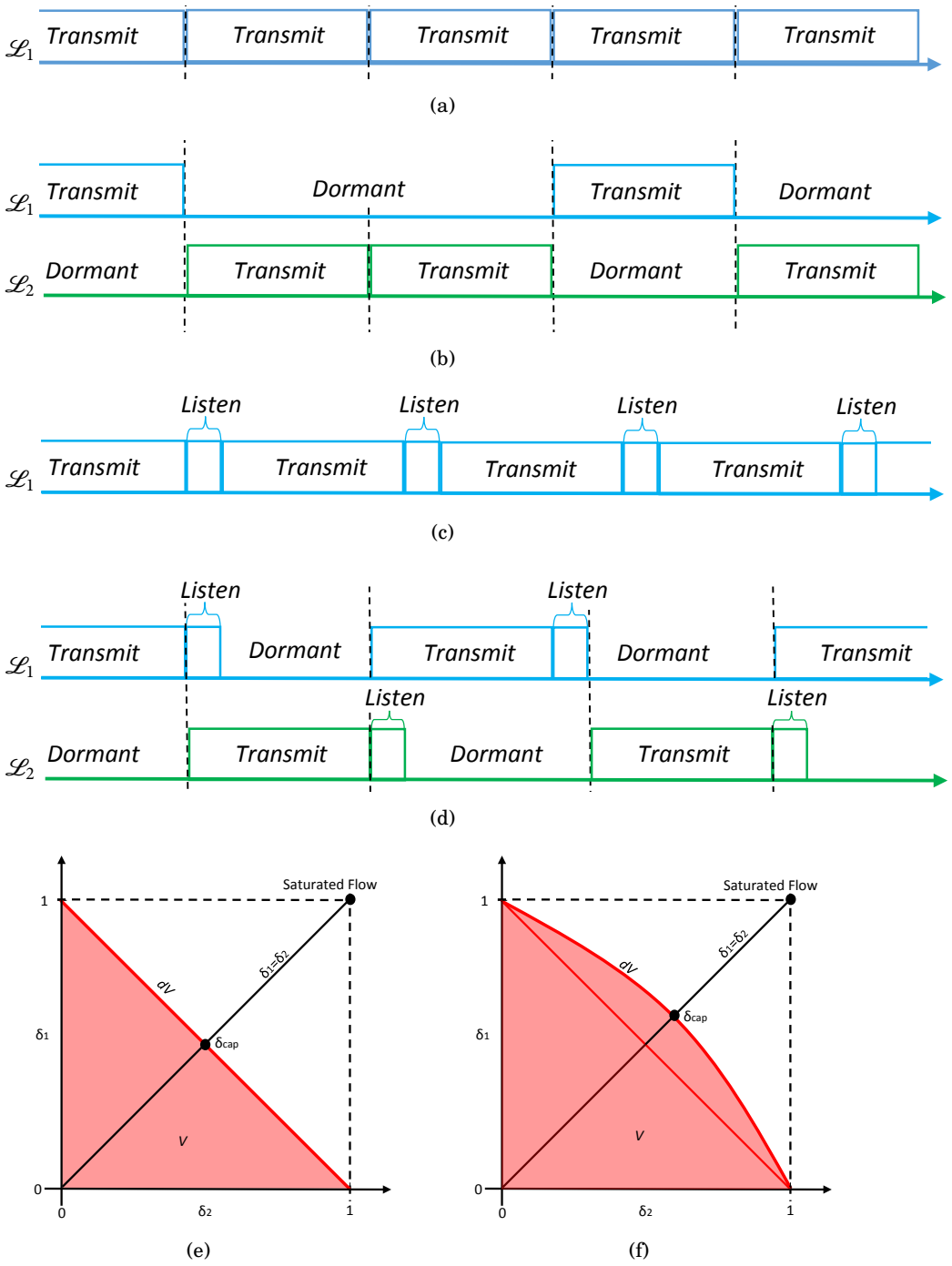


FIGURE 4.8. An illustration of why introducing the *Listen* state enables a convex capacity region. (a) Shows the four-state model with a single saturated demand transmitting alone. Time proceeds to the right. (b) Shows the four-state model with two demands sharing the channel without collisions. (c) Shows the five-state model with a single saturated demand transmitting alone. Note, after each *Transmit* state the model must wait in *Listen* before it can *Transmit* again. (d) Shows the five-state model with two demands sharing the channel without collisions. Now the *Transmit* of the second link can begin while the first is in *Listen*, hence capacity is increased. (e) and (f) show capacity region schematics for (b) and (d) respectively.

4.4.2 Fixed Parameters: *Transmit* and *Listen*

For all simulations in this thesis, a constant packet payload and therefore constant length of the *Transmit* state is assumed. The parameter is set to a constant 200 timesteps. Once a link enters the *Transmit* state, it will remain there for the duration of the 200 timesteps, even if a collision occurs during that time period. The collision will only be detected by the link at the end of the *Transmit* state.

Under ideal CSMA/CA conditions, a simple topology of two links within carrier sensing range is constructed in NS3 and within the conceptual simulation model for both the four-state and the five-state models, such that the results can be compared. With NS3 and both the four-state and five-state conceptual simulation models, a large ensemble of simulations are performed, with different demand intensity vectors for each of the two links (δ_1,δ_2). Each of the demands δ_1 and δ_2 is varied independently in 40 equal increments, resulting in $41^2 = 1,681$ simulations for each of the three set-ups. The duration of each simulation is set equivalent to the time needed to transmit $10,000$ packets without collisions. A capacity region is generated.

A CSMA/CA simulation under ideal conditions in NS3 produces a convex capacity region boundary, as per the findings of Laufer and Kleinrock [151]. Simulating the four-state conceptual simulation model produces a flat capacity region boundary. To match the convex capacity region shape of the NS3 simulation, the additional *Listen* state must be added to the four-state model to give the five-state model. (Note, to produce ideal conditions in the conceptual simulation model, *Countdown*, which is designed to introduce uncertainty into the model, is set to zero meaning it is not used. Further, as conditions are ideal, no collisions occur and so the model never enters *Backoff* — the *Countdown* and *Backoff* will be further explained in Sec. 4.4.3). Fig. 4.7 explains that due to the 802.11 protocols enforced backoff immediately after a successful transmission, capacity can be increased by introducing a second demand sharing the channel. When two demands are sharing the channel the one whose backoff timer decrements first will access the channel, hence the average time between any two packets is reduced compared to a single demand. This gain can be replicated in the five-state conceptual simulation model by introducing the *Listen* state to similarly force a waiting time between consecutive transmissions along the same link. Fig. 4.8 explains that a second demand can then *Transmit* whilst the first is in the *Listen* state, hence resulting in an increase in capacity when multiple demands share the channel and a convex capacity region.

Numerous iterations of the ensemble were performed increasing the fixed number of timesteps of the *Listen* parameter from zero in order to find the appropriate length to match the convexity of the capacity region produced by NS3. Fifteen timesteps were found to be the most appropriate length to replicate the convex shape. The capacity region boundaries for the four-state, five-state and NS3 CSMA/CA simulations under ideal conditions are plotted in Fig. 4.9. The ability of the five-state model to closely replicate NS3 is clear.

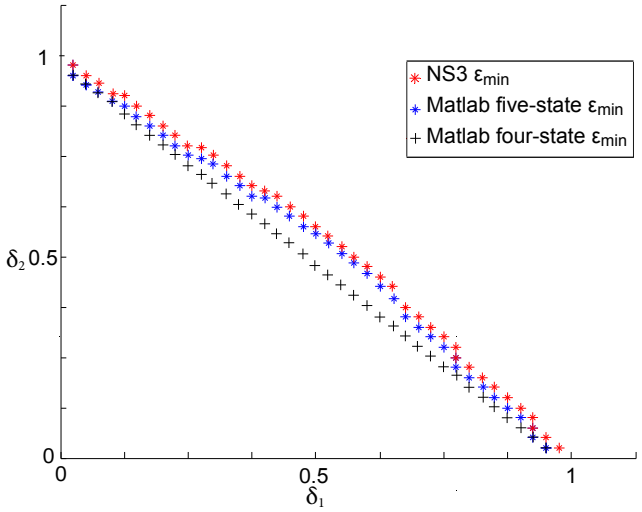


FIGURE 4.9. Validation of *Listen* parameter, comparison of conceptual simulation model with NS3. Capacity region plot for two competing links. Markers indicate the highest demand combinations that can be met by the system, i.e., the capacity region boundary.

4.4.3 Uniformly Distributed Parameters: *Backoff* and *Countdown*

The previous chapter discussed different interpretations of the backoff mechanism in the CSMA/CA modelling literature (Sec. 3.2.4 p.72). The CSMA/CA backoff's primary function is collision resolution, however, as Bianchi [29] explains, it is further related to the probability of a collision in the carrier sensing process (i.e., carrier sensing error ε). He defines a range for this probability, then throughout his analysis, he assumes ε at the lower bound (effectively no collisions). Laufer and Kleinrock [151] consider an ideal CSMA/CA with no collisions. Papers [8, 145, 250] argue these models overestimate goodput and do not adequately account for the cost of collisions even when considering ideal conditions. When operating in a PHY environment with some propagation loss and fading, it would be more realistic to consider a higher collision probability [129, 145]. When collisions are able to occur, the relatively short time spent in backoff is negligible to the longer time wasted by a collision.

Similar to the CSMA/CA protocol backoff, the conceptual simulation model *Backoff* state is designed to resolve contention between competing links following a collision. Fig. 4.10 illustrates how the *Backoff* state resolves a collision in the four-state conceptual model. States of the conceptual simulation model, however, separate collision resolution and control of ε. The *Backoff's* function is collision resolution and the *Countdown* state controls the level of ε. Both of these states are heuristically designed and have been refined through iterative development and validation of the conceptual simulation model.

The conceptual simulation model's *Backoff* state further differs from that of the CSMA/CA

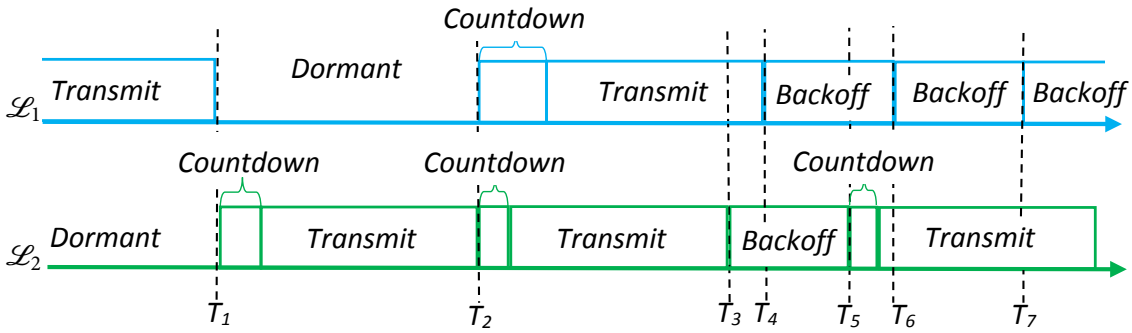


FIGURE 4.10. An illustration of why the *Countdown* state can cause collisions and how the *Backoff* state resolves collisions. Increasing the length of the *Countdown* will increase the chance of a collision. Time proceeds to the right. Key times are indicated. At T_1, $\mathscr{L}_1$ finishes *Transmit*. $\mathscr{L}_2$ senses the channel clear and moves to *Countdown* before *Transmit*. At T_2, $\mathscr{L}_2$ finishes *Transmit*. $\mathscr{L}_1$ senses the channel clear and moves to *Countdown* before *Transmit*. While $\mathscr{L}_1$ is in *Countdown*, $\mathscr{L}_2$ also senses the channel clear and moves to *Countdown* before *Transmit*. At T_3, $\mathscr{L}_2$ finishes *Transmit* and realises a collision has occurred. Hence, it proceeds to *Backoff*. At T_4, $\mathscr{L}_1$ finishes *Transmit* and realises a collision has occurred. Hence, it also proceeds to *Backoff*. At T_5, $\mathscr{L}_2$ finishes *Backoff*, senses the channel clear and moves to *Countdown* before *Transmit*. At T_6, $\mathscr{L}_1$ finishes *Backoff*, senses the channel busy and so must *Backoff* again. At T_7, again $\mathscr{L}_1$ finishes *Backoff*, senses the channel busy and so must *Backoff* again.

protocol backoff in a number of other ways. Firstly, for simplicity of implementation, the *Backoff* is only initiated on a link following a collision, and instead of increasing in size with each transmission attempt, is always chosen from the same uniform distribution. Several works show that in CSMA/CA, it is not necessarily exponentially distributed anyway [33, 39, 77, 93, 186, 266]. Further, the *Backoff* counter does not freeze when the node senses the channel busy, rather it continues to decrement and only senses the channel when it reaches zero. If then it senses the channel busy it selects a new *Backoff* from the same uniform distribution as before.

The *Countdown* is an artificial state introduced into the conceptual simulation methodology to add a level of error in the carrier sensing process. Each time the model enters the *Countdown* state its parameter is chosen from a uniform distribution. If the upper and lower bound of the *Countdown* are both set to zero, the simulator senses the channel perfectly. This means for transmitters within carrier sensing range of each other competing for channel access, collisions never occur. This is representative of the CSMA/CA process, as in NS3, when simulating under ideal conditions ($\varepsilon_{\min}$), where links sense the channel instantaneously and do not collide. Increasing the time period of the *Countdown* increases the delay between the transmitter sensing the channel is clear and beginning its transmission. This time, although still very short relative to the length of a transmission, allows time for another transmitter to sense the channel clear

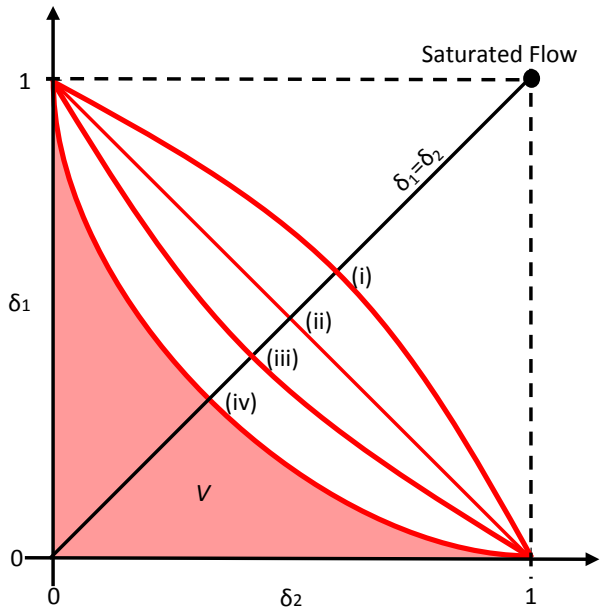


FIGURE 4.11. An illustration of how increased level of carrier sensing error (ε) increases the concavity of the capacity region. The change in boundary reflects a change in the level of error as follows. (i) Shows the convex boundary of CSMA/CA or the five-state model assuming ideal conditions (i.e., perfect PHY, zero propagation loss, instant carrier sensing and no collisions). (ii) Shows the flat boundary of the four-state model for ideal conditions. (iii) Shows a concave boundary for CSMA/CA and the four-state or five-state model when conditions are not ideal. Note the marginal gain in capacity of CSMA/CA from two demands sharing the channel is lost. Additional time spent resolving collisions makes small gains achievable by two demands sharing the channel redundant. (iv) Shows a greater concavity in the boundary due to a higher level of ε. As the CSMA/CA capacity region boundary becomes concave when ε is introduced, the *Listen* state is no longer required in the five-state model and the four-state model can be used to represent CSMA/CA.

and similarly begin its transmission, therefore leading to a collision. Increasing the *Countdown* increases the probability of a collision between transmitters in competition for channel access. Fig. 4.10 illustrates how the introduction of the *Countdown* may lead to a collision, with multiple transmitters sensing the channel clear in quick succession. Fig. 4.11 illustrates how increasing the level of ε (hence increasing *Countdown*) impacts the capacity region for two demands sharing channel access.

Three experiments for validating the parameters of the conceptual simulation model against results generated in NS3 are provided in Secs. 4.4.4, 4.4.5 and 4.4.6. The level of ε in NS3 is manipulated by adjusting the propagation loss. Tab. 4.2 presents the bounds of the uniform distribution from which the *Countdown* is selected, that are required to replicate simulation results produced in NS3 for ε_{min}, ε_{mid} and ε_{max}. For all simulation results displayed in this

Table 4.2: *Countdown* parameters required to replicate specified level of carrier sensing error (i.e., the probability an RTS generated does not lead to successful transmission of CTS, data packet and ACK) in CSMA/CA networks.

	Carrier sensing error	NS3 Propagation Loss Between Link Pairs (db)	$Countdown_{\min}$	$Countdown_{\max}$
$\varepsilon_{\min}$	0	0	0	0
ε_{mid}	0.03	50	0	2
$\varepsilon_{\max}$	0.06	54	0	8

chapter, $Backoff_{\min} = 16$ and $Backoff_{\max} = 64$ are used for both the five-state and four-state models. Of course when considering $\varepsilon_{\min}$ (i.e., $Countdown_{\max} = 0$), collisions do not occur and the model never enters *Backoff*.

Further to much informal exploration and several iterations of the three experiments outlined in the following sections, the fixed bounds of the uniform distribution of the *Backoff* parameter were chosen as follows. Considering $\varepsilon_{\max}$ (i.e., $Countdown_{\max} = 8$ in the four-state model), for all feasible combinations of the upper and lower bound from the range $Backoff_{\min} = 2,4,...,32$ and $Backoff_{\max} = 2,4,...,80$, a large ensemble of simulations were conducted. Vectors of demand intensity (δ_1, δ_2) were simulated from 0 to 1 in 40 equal increments for two links in competition for channel access, hence $41^2 = 1,681$ simulations for each pair of *Backoff* parameters. An equivalent simulation of 1,681 demand combinations to two links in competition for channel access was conducted in NS3. The capacity region was derived for all. For each pair of *Backoff* parameters used in the conceptual simulation model, the 1,681 demand combinations simulated were individually compared to the equivalent NS3 simulation results. The individual demand combinations simulated were compared to determine whether they had been judged in or outside the capacity region by both NS3 and the conceptual simulation model, or whether, NS3 and the conceptual simulation model had judged them differently. A figure was recorded determining the number of simulations that did not correlate from NS3 to the conceptual simulation model, hence the *Backoff* parameters for the conceptual simulation model with the lowest number of simulation not correlating to NS3 could be identified, these were $Backoff_{\min} = 16$ and $Backoff_{\max} = 64$ (note, small changes to either of these bounds make little difference to the model's performance). As is shown in the three experiments that follow, results generated in NS3 can be closely replicated using the conceptual simulation model with this *Backoff* range for each level of ε. The boundary of the capacity region for two links in competition, as in the described experiment, is shown in Fig. 4.12, overlaying the conceptual simulation model and NS3.

For each of the three experiments that follow six scenarios are simulated: three simulations with NS3, one at each of the three levels of ε, and three with the conceptual simulation model in Matlab, each to replicate one of the three levels of ε. For the ideal CSMA/CA case $\varepsilon_{\min}$, the five-state conceptual simulation model is used to replicate NS3, and for the other two levels of

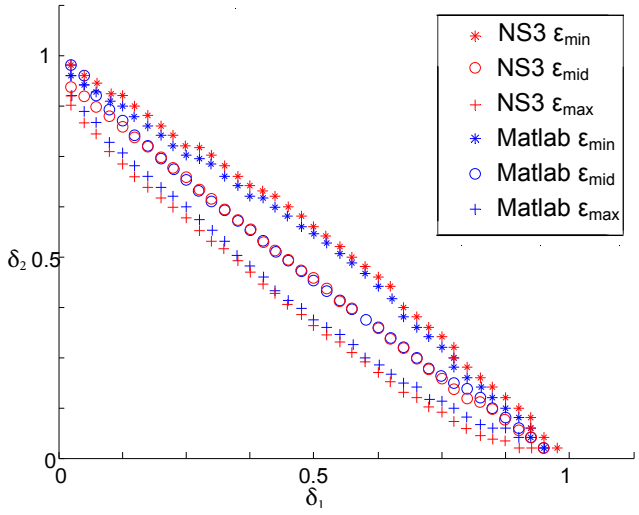


FIGURE 4.12. Validation of the conceptual simulation model by comparison with NS3. Capacity region boundary plot for two competing links at three levels of carrier sensing error (ε). Markers indicate the highest demand combinations that can be met by the system, i.e., the capacity region boundary.

error ε_{mid} and ε_{max}, the four-state model is used.

4.4.4 Validation Experiment 1: 2 and 3 Links

This experiment considers networks of just two and three links in direct competition for channel access. For each of the six scenarios, a large ensemble of simulations are performed, each with a different demand intensity vector (δ_1,δ_2,...,δ_3). Each of the demands is varied independently from 0 to 1 in 40 equal increments, resulting in $41^2 = 1,681$ simulations for two link cases and $41^3 = 68,921$ simulations for three link cases. The duration of each simulation is set at the time needed to transmit 10,000 packets without collisions.

Fig. 4.12 plots the boundary of the capacity region of each of the two link examples. Demand combinations with a total less than the plotted boundary (i.e., below the line of the boundary) are within the capacity region. Total demand combinations above the plotted boundary are outside the capacity region, i.e., one or more queues grow unbounded.

It can be observed from the results of the CSMA/CA simulations in NS3 (Fig. 4.12) that at ε_{min}, the capacity region is convex. This is equivalent to the findings of Laufer and Kleinrock [151]. Further, as Jindal and Psounis [129] suggested, when a non-zero amount of ε is introduced, results from CSMA/CA simulations in NS3 show that the capacity region is in fact concave. The additional time the CSMA/CA protocol spends contesting the channel, and in backoff following an RTS collision, removes any potential gain in capacity from the introduction of a second demand. Fig. 4.12 illustrates how these capacity regions can be replicated by the conceptual simulation

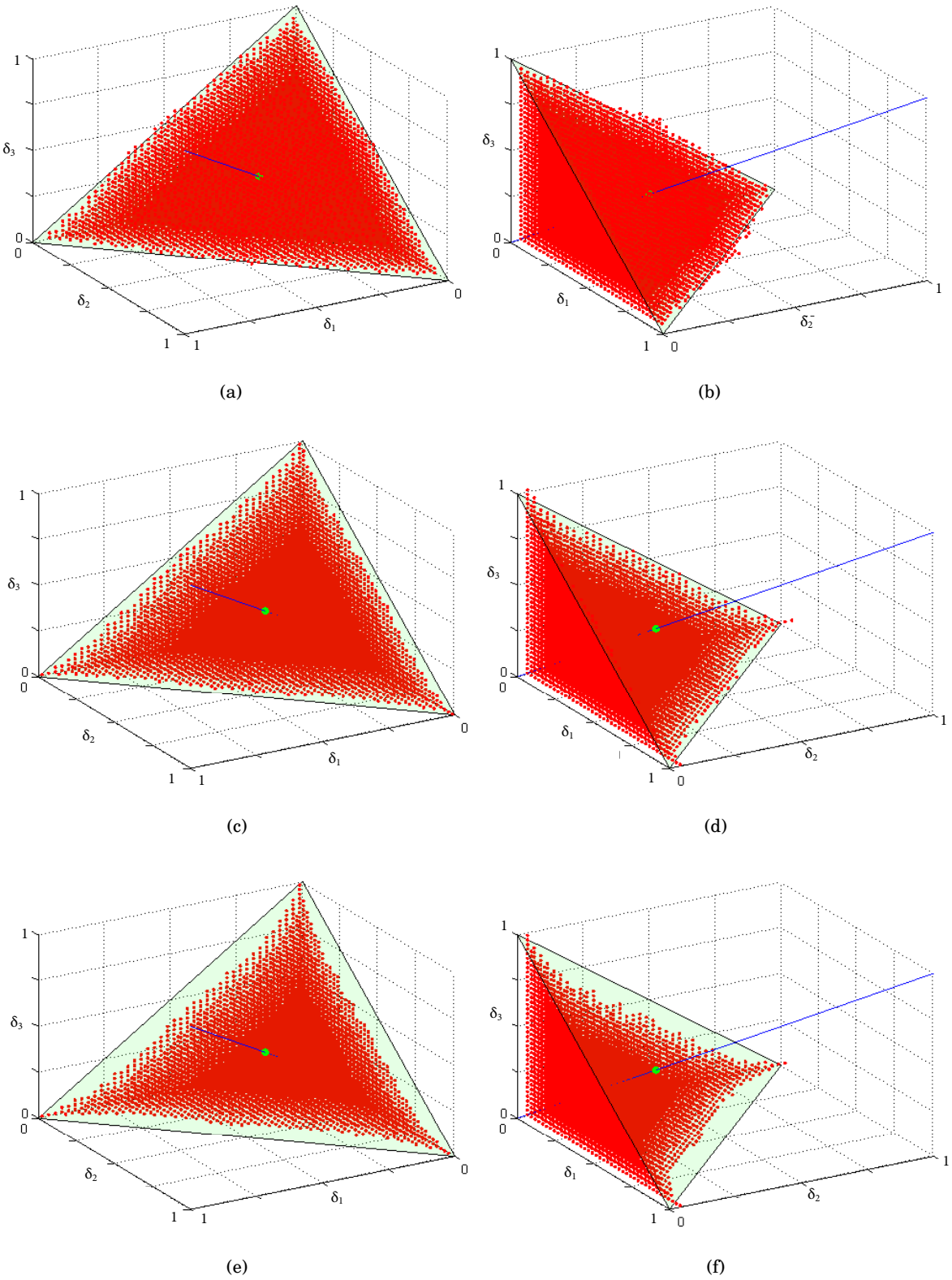
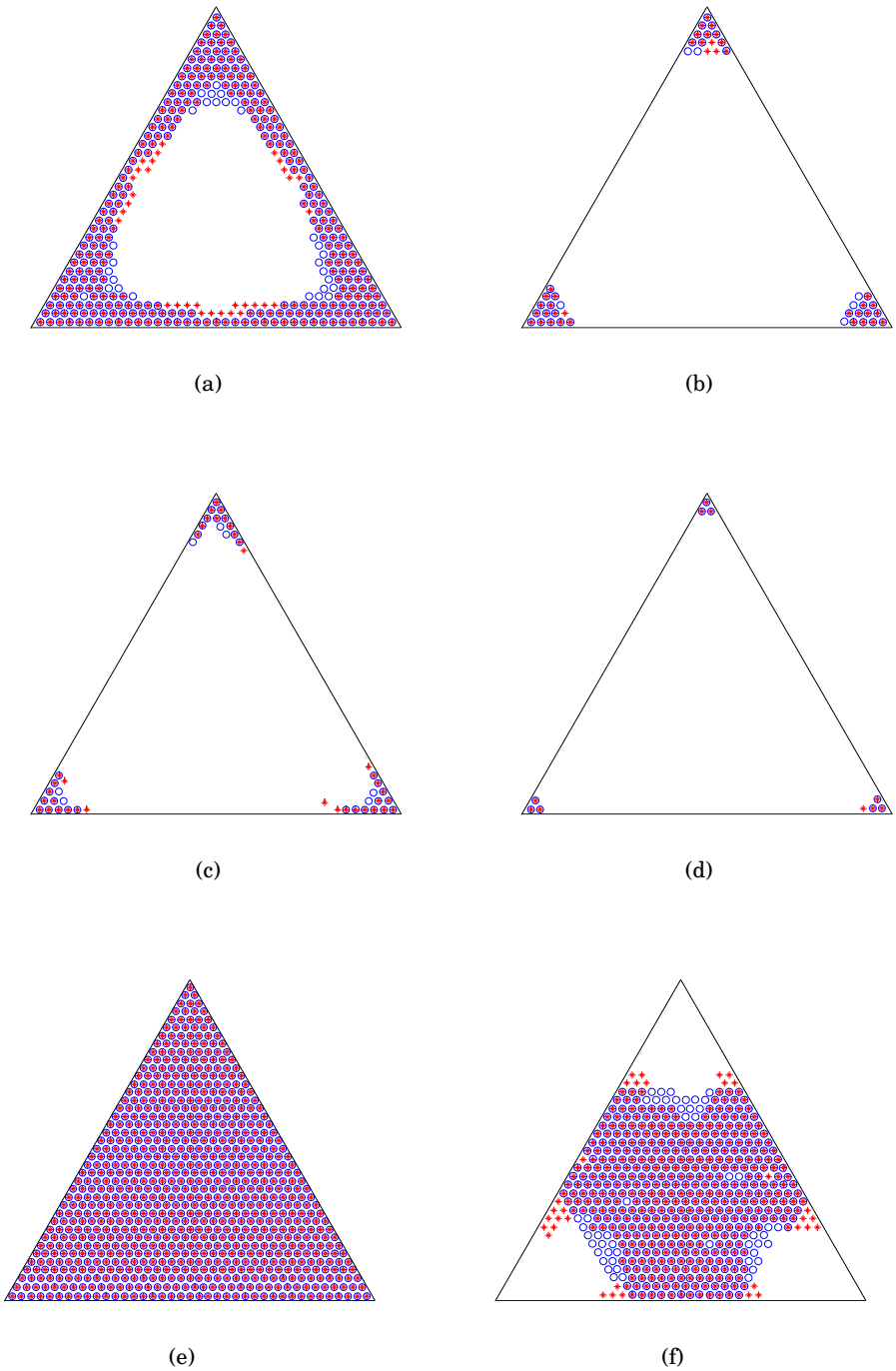


FIGURE 4.13. Capacity region plots for three competing links, generated using the conceptual simulation model. Red markers indicate demand combinations within the capacity region. (a) & (b) $\varepsilon_{\min}$. (c) & (d) ε_{mid}. (e) & (f) $\varepsilon_{\max}$.



(a) (b) (c) (d) (e) (f)

FIGURE 4.14. Triangular sections $\delta := \delta_1 + \delta_2 + \delta_3 = \text{const.}$ (see Fig. 4.13). Red markers and blue circles indicate points inside the capacity region for the conceptual simulation model and NS3 respectively. The sections show total demand and the level of carrier sensing error as follows: (a) $\delta = 2.85$, ε_{mid}, (b) $\delta = 2.85$, ε_{max}, (c) $\delta = 3.00$, ε_{mid}, (d) $\delta = 3.00$, ε_{max}, (e) $\delta = 3.00$, ε_{min}, (f) $\delta = 3.15$, ε_{min}

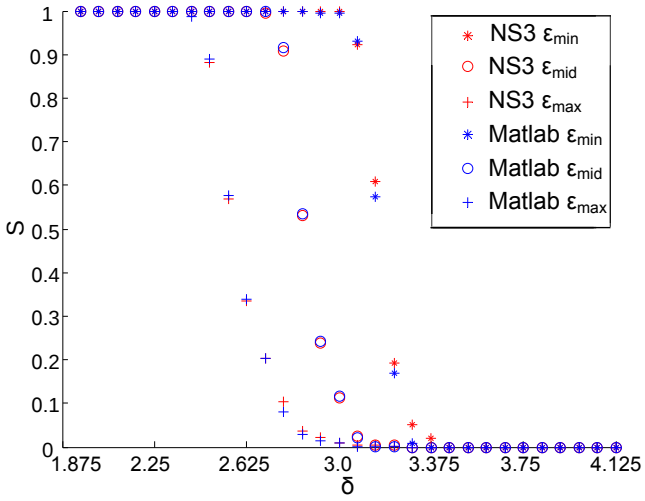


FIGURE 4.15. Validation of conceptual simulation model by comparison with NS3. Proportions S of the triangular section within the capacity region, as a function of total demand intensity $\delta := \delta_1 + \delta_2 + \delta_3$, compared across three levels of carrier sensing error ε.

model.

Fig. 4.13 shows two perspectives of the three-dimensional capacity region plot for three links in competition generated using the conceptual simulation model. Red dots indicate a point is inside the capacity region. In Figs. 4.13(a) and 4.13(b), from simulation of $\varepsilon_{\min}$, the slight convexity of the capacity region can be observed. In Figs. 4.13(c) and 4.13(d), when carrier sensing error is introduced, a slight concavity is visible and this concavity is more apparent for Figs. 4.13(e) and 4.13(f) where the level of carrier sensing error is higher.

To better analyse Fig. 4.13 and compare the conceptual simulation model with NS3, consider total demand intensity $\delta := \delta_1 + \delta_2 + \delta_3$. Then sections $\delta = \text{const.}$ describe triangular cross-sections through Fig. 4.13. Examples of different triangular sections at different values of δ, overlaying results from the conceptual simulation model with NS3, are shown in Fig. 4.14. The dependence of the proportion of the corresponding triangular area that is within the capacity region S on δ may then be studied and compared across the six scenarios: see Fig. 4.15. Although there are small discrepancies in the classification of individual demand combinations, it is clear that by performing a large ensemble of simulations, the statistics of the capacity region generated in NS3 can be closely replicated with the conceptual simulation model.

The results shown in Figs. 4.12 and 4.15 illustrate the significance of the range of ε derived in [29]. When considering $\varepsilon_{\min}$, the shape of the capacity region is convex, and hence introducing new demands into a system increases the total system capacity. When considering ε_{mid} or $\varepsilon_{\max}$, the shape of the capacity region is concave, indicating a capacity penalty when a new demand is introduced.

Increasing ε will increase the concavity of the capacity region and reduce the saturation goodput. Paper [145] showed how the capacity region became increasingly concave with increased propagation delay (and hence ε) between a pair of nodes in competition for channel access.

4.4.5 Validation Experiment 2: Service and Latency Distribution

Consider three links in competition for channel access: two links ($\mathscr{L}_1$ and $\mathscr{L}_2$) have fixed demand intensity $\delta = 1/3$, the third link's ($\mathscr{L}_3$) δ is varied from 0 to 1/3 in 44 equal increments, resulting in 45 simulations. The duration of each simulation is set at the time needed to transmit 100,000 packets without collisions.

Relative to the moment each packet arrives in the queue, the total proportion of packets served by $\mathscr{L}_3$ in the comparable time for $1,2,..,4$ consecutive packets to be transmitted in a clear channel is recorded. Fig. 4.16 (a), (c) and (e) plot this value against δ.

The latency distribution of $\mathscr{L}_3$ for all packets served at each value of δ can be observed, overlaying NS3 and conceptual simulation model results. Figs. 4.16 parts (b), (d) and (f) plot sections of the latency distribution.

For Fig. 4.16 parts (c) and (e), the similarity between the conceptual simulation model and NS3 results is clear. Part (a) shows a slight divergence between the conceptual simulation model and NS3 for middle range values of δ, however, the two converge at higher δ (i.e., intersection with the x-axis as the system approaches saturation). The progressive left shift of the plot (a), (c), (e) and their points of intersection with the x-axis indicate the increasing loss of efficiency of the more error-prone systems. Similarly, each latency plot shows similar behaviour for the two simulations.

4.4.6 Validation Experiment 3: 1-20 Links

This experiment considered networks of $n = 1,2,...,20$ links. Demand intensity δ is applied equally to all links, increasing from 0 to 1 in 1000 equal increments, resulting in 1,001 simulations for each network. Hence, 20,020 individual simulations are conducted for the six scenarios. The duration of each simulation is set for the time needed to transmit 100,000 packets without collisions. The traffic demand δ_{cap} at which the network becomes saturated is identified.

Fig. 4.17 shows the total traffic required to saturate the network. NS3 results are replicated with the conceptual simulation model for the three levels of error investigated. Here, the significance of the RTS collision probability range derived by Bianchi [29] is apparent. When considering the lower bound of error in the carrier sensing process (ε_{min}), the addition of new demands appears to increase the saturation point of the network. When considering higher values of ε (ε_{max} or ε_{mid}), the saturation point is reduced with the addition of new demands.

Plotting the results of experiments for ε_{min} in Fig. 4.17 supports results in Bianchi's work [29] showing the potential of additional demands in a CSMA/CA system to increase saturation goodput. Plotting the results for ε_{max} more closely resembles the findings of Tinnirello, Bianchi

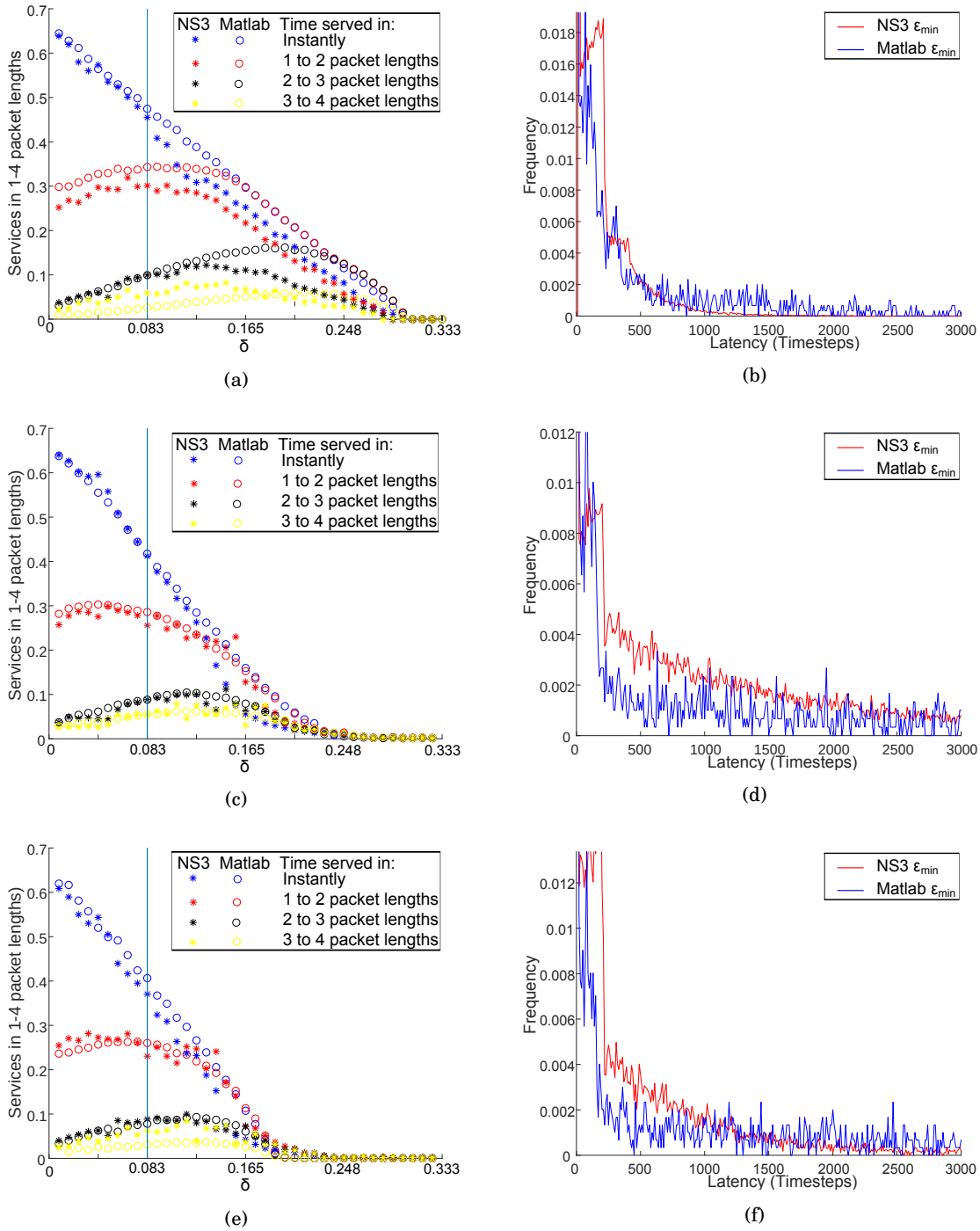


FIGURE 4.16. Each pair of plots shows on the left; the proportion of packets served; instantly (blue), in one to two packet lengths (red), in two to three packet lengths (black) and in three to four packet lengths (yellow). Comparing NS3 (markers) to the conceptual simulation model in Matlab (circles). And on the right: the latency distribution for $\delta = 0.083$ (vertical blue line in corresponding plot). The carrier sensing error is set as follows: (a) and (b) $\varepsilon_{\min}$, (c) and (d) ε_{mid}, (e) and (f) $\varepsilon_{\max}$.

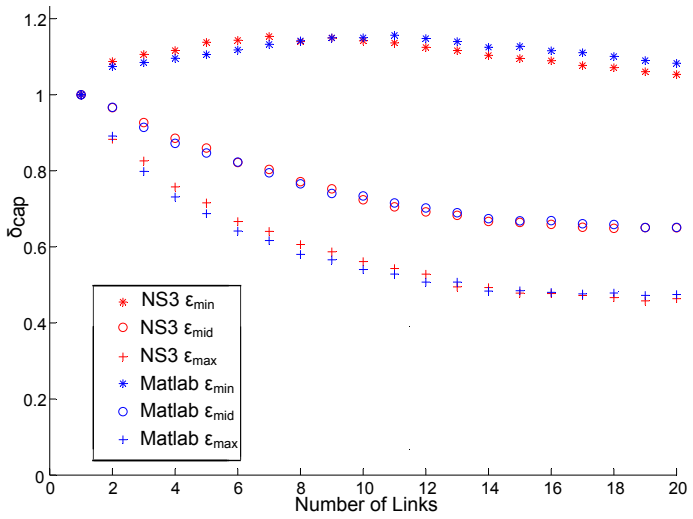


FIGURE 4.17. Validation for 1 to 20 competing links of equal demands by comparison of the conceptual simulation model with NS3. Plot shows the saturation points δ_{cap} for 1 to 20 competing links of equal demands.

Table 4.3: Calibrated conceptual simulation model parameters.

Parameter	Timesteps
Dormant	-
Transmit	200
Listen (five-state model Only)	15
$Backoff_{\text{min}}$	16
$Backoff_{\text{max}}$	64
All levels ε — $Countdown_{\text{min}}$	0
ε_{min} — $Countdown_{\text{max}}$	0
ε_{mid} — $Countdown_{\text{max}}$	2
ε_{max} — $Countdown_{\text{max}}$	8

and Xiao [250] or Jindal and Psounis [129] who refined Bianchi's original model and found, under a more realistic interpretation of CSMA/CA (i.e., not ideal) additional demands will decrease goodput saturation point (see discussion of the work of [250], [129] and [29] in Sec. 3.2.4 p.72).

As Fig. 4.17 indicates, the simulations considering three levels of carrier sensing error generated in NS3 can be closely represented by the conceptual simulation model.

4.5 Conclusion

This chapter contributes to knowledge in the following ways:

CTK 4.1 Presenting a valid conceptual simulation modelling methodology capable of representing wireless CSMA/CA networks.

CTK 4.2 Identifying that when considering zero carrier sensing error the capacity region is convex, and hence introducing new demands into a system increases the total system capacity. By increasing carrier sensing error (only slightly), the capacity region becomes concave, indicating a capacity penalty when a new demand is introduced. The concavity increases with increasing carrier sensing error.

This chapter proposes and validates a modelling methodology for investigating wireless CSMA/CA networks. Small discrepancies in classification of individual demand combinations and differences in the internal service and latency distributions can be seen, in validation experiments 1 and 2, between the conceptual simulation modelling methodology and NS3. However, it is clear that the conceptual simulation modelling methodology is able to closely replicate the capacity region produced via NS3 simulation. Tab. 4.3 summarises the model parameters. The chapters that follow extend this methodology, fitting it to situations which can provide novel insights into proposed next-generation technologies.

4.A Appendix

4.A.1 Meta Code

The core meta code for the five-state conceptual simulation model (see Fig 4.1(a)) is provided as follows. With each timestep all links of the network cycle through this loop. For the four-state model (see Fig 4.1(b)) remove the *Listen* state such that when a successful transmission is completed the link returns to the *Dormant* state:

```
if isDormant(j)
    % if there is a packet waiting in the queue, check the channel, if clear, enter Countdown to
transmit
    if isNodeClearToTransmit(j) && queue(j)>0
      isDormant(j)=false;
      isCountdown(j)=true;
      counter(j)= random value selected from Countdown uniform distribution;
end
elseif isTransmit(j)
      counter(j) = counter(j)-1;
      % Check if a collision has occurred
  if  isClashed(j)
      isClashed(j) = sum colliding transmissions ;
  end
  if counter(j)==0
  if isClashed(j)
      % packet has collided during transmission
      isTransmit(j)=false;
      isBackoff(j)=true;
      counter(j) = random value selected from Backoff uniform distribution;
      % record failed transmission
      failedTransmissionsTotal(j) = failedTransmissionsTotal(j) + 1;
  else
      % successful transmission, record time packet arrived in queue, record transmission time,
record latency.
      LatencyTransmission(i,j)= Service Time - Arrival Time;
      % remove packet from queue
      queue(j)=queue(j)-1;
      isTransmit(j)=false;
      isListen(j)=true;
      counter(j)= set counter to Listen param;
```

```
      goodTransmissionsTotal (j) = goodTransmissionsTotal(j) + 1;
      end
   end
   end
elseif isListen(j)
   counter(j) = counter(j)-1;
   if counter(j)==0
      isListen(j)=false;
      isDormant(j)=true;
   end
elseif isBackoff(j)
   counter(j) = counter(j)-1;
   if counter(j)==0
      if isNodeClearToTransmit(j)
         isBackoff(j)=false;
         isCountdown(j)=true;
         counter(j)= random value selected from Coundown uniform distribution;
      else
      counter(j) = random value selected from Backoff uniform distribution;
      end
   end
elseif isCountdown(j)
   counter(j) = counter(j)-1;
   if counter(j)==0
      isCountdown(j)=false;
      isTransmit(j)=true;
      isClashed(j)=false;
      counter(j)= value set to length of Transmit;
   end
end
```

CHAPTER 5

WIRELESS NETWORK MAC LAYER PERFORMANCE EVALUATION WITH FULL-DUPLEX CAPABLE NODES

This chapter is based on: W. Jones, R. E. Wilson, M. Sooriyabandara, and A. Doufexi, "Wireless Network MAC Layer Performance Evaluation with Full-Duplex Capable Nodes," in Proceedings of the 12th ACM Symposium on QoS and Security for Wireless and Mobile Networks - Q2SWinet'16, 2016, pp. 111-118.

Through developments in self-interference cancellation technologies, the potential to significantly increase link capacity in wireless networks has emerged by enabling nodes to transmit and receive simultaneously in full-duplex [58, 82, 124, 206]. This development suggests the implementation of full-duplex nodes to be a favourable addition, dramatically increasing performance for next-generation communications networks. Whilst evidently capable of improving the capacity of individual links, the gain from a full-duplex system implemented in a network has been questioned and analysis has shown that the capacity gain is limited due to aggregate interference [102, 267].

This chapter provides two example studies which demonstrate the potential impact on network performance that full-duplex can provide. The studies extend the conceptual simulation methodology outlined in Chapter 4.

The first study in Sec. 5.2 demonstrates the potential capacity gain that full-duplex can provide applied in various forms to a highly simplified mesh network set-up. The study shows that full-duplex alone can increase the capacity of a network, however, when full-duplex is combined with appropriate interference management, the gain is much more significant. Existing work hints at this result but this study is the first to quantify the gains. The majority of the capacity gain was shown to occur at asymmetric demand combinations which would not have

been observed only considering equal or saturated demands as is traditionally the case.

The second study, in Sec. 5.3, focuses on two common issues in current communications networks, namely bottlenecks and hidden nodes, and demonstrates the potential performance gains that full-duplex nodes can offer via a simple example of two clients and an access point. The study shows that introducing full-duplex access points alone mitigates against the problem of bottlenecks, reduces the impact of hidden nodes, and can increase the capacity of a network. When full-duplex access points are able to work with full-duplex clients, the capacity gain is much more significant, however, it is shown that much of this capacity gain occurs at uneven demand combinations. When the demand to all nodes is equally high, the introduction of full-duplex capability to clients is shown to increase the number of transmission attempts resulting in a significantly increased number of collisions and reduced network performance. Further, it is observed that at low traffic levels, a full-duplex access point may improve goodput by simply transmitting a busy tone to silence other transmissions whilst it receives, mitigating against the hidden node problem.

Physical layer management in relation to full-duplex is still an open research topic. For example, interference management techniques such as beam-forming, sectorisation and directional diversity [82, 102, 267] are under active investigation and their implementation in some form may be required to achieve the full-duplex scenarios described. It is acknowledged that there are challenges presented in the topologies of the two studies in this chapter (see Fig. 5.1 p.131 & Fig. 5.7 p.137). These are: firstly self-interference at the base station caused by simultaneous transmission and reception (Fig. 5.2 p.132 & Fig. 5.8 p.139 both bidirectional and unidirectional); secondly interference from the uplink transmitter at the downlink receiver (Figs. 5.2 & 5.8 unidirectional). In the first case the challenge is how to communicate in the presence of high power interference and in the second the challenge is how to communicate in the presence of a commensurate-power interference [82]. The literature on interference channels and managing interference is extensive, but a definitive understanding of managing full-duplex interference is still under development. Acknowledging that both challenges are significant and need to be addressed, in the studies that follow a physical layer capable of managing this interference is assumed possible and the study focuses on the MAC layer interactions. Further details of the assumptions applied are given in each example.

The overall purpose of this chapter is to evaluate the impact on performance of introducing full-duplex nodes into communications networks. The contribution of this chapter is a guide indicating the impact on capacity gain and network performance full-duplex can provide and a methodology for exploring this.

5.1 Background

Wireless communication signals decay quickly with distance, and hence in a wireless network, the transmit signal powers are much higher than received signal powers. Because of this, it has long been thought that a wireless node cannot transmit and receive on the same frequency at the same time, as the high power transmit signal would result in unavoidable self-interference at the receiver [152]. In current radio systems, simultaneous transmission is achieved by simply avoiding this problem, dividing transmit and receive across different channels. First in the 1980s [51], and further, more recently [13, 26, 58, 78, 140, 221], research has proposed and demonstrated systems allowing simultaneous transmission and reception at the same frequency, commonly known as full-duplex wireless. This potentially has significant benefits for spectral efficiency, theoretically providing double the capacity without increasing bandwidth.

For a node to operate in full-duplex, self-interference must be minimised. Any residual self-interference will add to the background noise, reducing the channel capacity. To achieve the theoretical doubling of capacity, perfect isolation (i.e., completely free of any interference or noise) would be required. Perfect transmitter-receiver isolation is not practically feasible, however, provided the isolation is high enough such that the residual self-interference is a significant margin below the existing background noise at the receiver, then the impact on the receive performance will be negligible [152].

Existing designs [13, 26, 31, 51, 58, 78, 79, 155], use combinations of analogue and digital cancellation along with antenna-based suppression to achieve the required level of isolation. Digital cancellation [78, 79, 89, 124, 140], typically the easiest to implement, although effective, cannot achieve the significant isolation required by itself. Analogue cancellation can [26, 58, 78, 136, 155, 221, 228], making it a requirement for full-duplex systems [26]. Hardware limitations and system performance inefficiencies have been explored in depth [30, 70, 213, 220]. Further network capacity and refinements to MAC protocols for full-duplex networks have been investigated [20, 37, 124, 134, 221, 222].

The work presented in this chapter supports the existing work seeking to identify the performance gain that full-duplex can offer communications networks in the MAC layer [52, 58, 82, 102, 124, 160, 206, 265, 267, 281]. In contrast to the approach of [52, 265, 267], who provide statistical data for large networks, this study presents detailed examination of simple network topologies. Existing studies use probabilistic arguments to derive the throughput in saturated conditions for networks of different sizes comparing existing Half-Duplex Medium Access Control (HD-MAC) with proposed new Full-Duplex Medium Access Control (FD-MAC) protocols [52, 160, 281]. Performance statistics, such as the likelihood of a successful packet transmission, are derived based on the concentration of nodes in the network. The performance benefit of full-duplex is then expressed in terms of throughput per unit area. Like [265], in both studies, a simplified network interference model is presented and the conceptual simulation model, presented in Chapter 4, applied. The novel approach of the conceptual simulation model

allows the study to scan over a large parameter space and identify the capacity region [252] in a way not computationally tractable with normative simulation. This approach enables the source of capacity gains to be identified.

Papers [102] and [267] suggest that the problematic aggregate interference in full-duplex networks will reduce the potential benefits. However, they suggest this could be mitigated with appropriate interference management techniques such as beamforming and sectorisation. The first study presented in this chapter aims to quantify the gain achievable from full-duplex in a wireless mesh network. Interference management techniques are considered, analysing simplified models of these also, quantifying the performance impact of these combined with full-duplex.

Past work exploring the potential performance gain of full-duplex includes [52] which proposes a FD-MAC protocol and uses a probabilistic argument to derive the throughput in saturated conditions for networks of different sizes. The same methodology is demonstrated with HD-MAC and the gain from full-duplex derived. Paper [267] used a similar approach and identifies transmission capacity in relation to network density based on the probability of a node accessing the channel. The gain of full-duplex is analysed in terms of average throughput per area in paper [265]. The authors apply a Poisson bipolar model [16] which inherits the simple interference structure of the classic Gupta and Kumar model [107].

Paper [276] applies a similar simplified interference model and identifies specific conditions under which full-duplex throughput gain is achieved. The paper compares full-duplex with other techniques where two RF chains can operate simultaneously, i.e., MIMO. The paper highlights the impact of network topology on the effectiveness of these techniques.

When a WLAN network is saturated, a bottleneck occurs at the access point. This occurs because the access point is subject to the same protocol, and thus has the same probability of gaining access to the channel for transmission as each client while serving more demands. It is suggested that full-duplex can unblock this bottleneck, because of the ability of a full-duplex access point to transmit simultaneously whenever it is receiving a packet from a client [124]. Throughout this chapter, this is referred as the 'bottleneck problem'.

Further in WLANs, the hidden node problem occurs when a half-duplex receiver cannot notify other nodes in the network that it is currently receiving a signal and thus other nodes attempt to transmit to the receiver (see Chapter 3 p.58). Implementing full-duplex might mitigate against the hidden node problem because of the additional ability of the receiving node to simultaneously transmit. This transmission suppresses nearby nodes thus eliminating the hidden node problem.

In the second study, a simple network model which captures the 'bottleneck' and 'hidden node' problems is presented. Implementing full-duplex in the model is investigated and the performance gains in the network are quantified.

Several works recognise the potential of full-duplex nodes to increase network capacity, reduce the effects of the 'bottleneck problem' or the 'hidden node problem', such as [219, 273], and further others begin to analyse the impact [124, 221]. Paper [124] implements a physical test

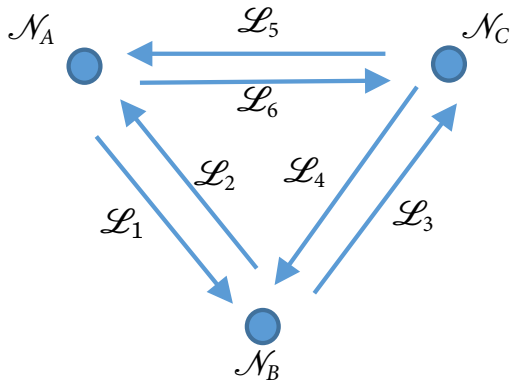


FIGURE 5.1. Simple three node mesh network. Each pair of the three nodes $\mathcal{N}_A$, $\mathcal{N}_B$ and $\mathcal{N}_C$ is connected by two of the six directional links $\mathcal{L}_1$, $\mathcal{L}_2$, ..., $\mathcal{L}_6$.

of two client devices hidden from each other attempting to communicate with an access point. They compare the system performance using a half-duplex and full-duplex access point and find significant improvement from the full-duplex access point. This research furthers this analysis using a simulation approach, however focusing the study on MAC layer interactions. Comparing simulations using the conceptual simulation model (see Chapter 4), using HD-MAC based on 802.11 and FD-MAC based on those proposed in [52, 160, 281] (some additional timesteps are added to the conceptual simulation model to reflect the longer protocol overhead), it is possible to quantify the resulting performance increase for the simple network setups when full-duplex is introduced.

The following Secs. 5.2 and 5.3 describe the two case studies, with results and analysis presented for each individually. Sec. 5.4 then discusses the findings from both studies and Sec. 5.5 draws conclusions.

5.2 Study 1: An Investigation into the Capacity Gain of Full-Duplex MAC in Wireless Mesh Networks

5.2.1 Simplified Mesh Network Setup

This study considers a three-node mesh network where all nodes can communicate directly with each other via a single hop, see Fig. 5.1. Each node ($\mathcal{N}_A, \mathcal{N}_B, \mathcal{N}_C$) can transmit to either of the other two nodes directly and hence there are six possible links (i.e., transmissions) that are denoted $\mathcal{L}_1, \mathcal{L}_2, ..., \mathcal{L}_6$.

It is supposed, further, that there is no routing and that each packet has a journey that consists of a single hop. It follows that one may prescribe demand (number of packets per unit time) on the network in terms of requested flows $d_1, d_2, ..., d_6$ for each of the six links respectively.

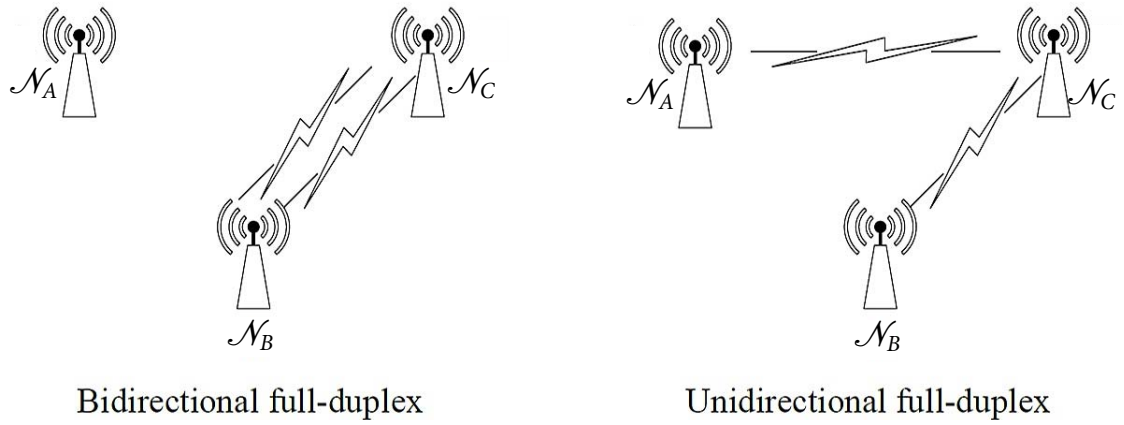


FIGURE 5.2. Bidirectional (3 possible variations) and unidirectional (6 possible variations) full-duplex for the simple mesh topology.

However, to simplify the computational analysis, the study is restricted to the special symmetric case $d_1 = d_6 =: d_A$, $d_2 = d_3 =: d_B$, and $d_4 = d_5 =: d_C$, so that the study need only scan through the three-dimensional demand space (d_A, d_B, d_C). Further, no routing is considered so that each packet reaches its destination via a single hop transmission.

Thus (for example) $\mathcal{N}_A$ is assumed to have a single queue of packets awaiting transmission, with Poisson arrivals of rate $2d_A$, and the queue will consist on average of an equal mix of packets bound for $\mathcal{N}_B$ and $\mathcal{N}_C$ on either $\mathcal{L}_1$ or $\mathcal{L}_6$. Further, the study supposes that there is no active queue management: packets are served in order of arrival, irrespective of their destination. Hence (for example) at $\mathcal{N}_A$, packets bound for $\mathcal{N}_C$ may be blocked by packets bound for $\mathcal{N}_B$ if link $\mathcal{L}_1$ is prone to collisions, even if link $\mathcal{L}_6$ is relatively collision-free.

Finally, the study supposes that signal strength is such that all nodes can hear each others' transmissions unless interference management techniques are in use. Thus the study focusses on the limitations of full-duplex due purely to direct interference, without any 'hidden node' effects [126] that one would expect to further limit capacity in larger mesh networks.

5.2.2 Operational Scenarios

One may now consider collisions between the various links within the network and how they might be mitigated by full-duplex techniques. The study will consider four operational scenarios: 1. Half-Duplex (HD); 2. (standard bidirectional - see Fig. 5.2) Full-Duplex (FD); and 3./4. two variants of Full-Duplex that employ interference management techniques (enabling unidirectional FD - see Fig. 5.2).

Throughout, the study supposes a simple deterministic binary interference model, so that if two incompatible links transmit simultaneously, there will be a collision with probability one. Thus a six-by-six collision matrix $\mathbf{E}$ is used (see Fig. 5.4) to describe each operational scenario,

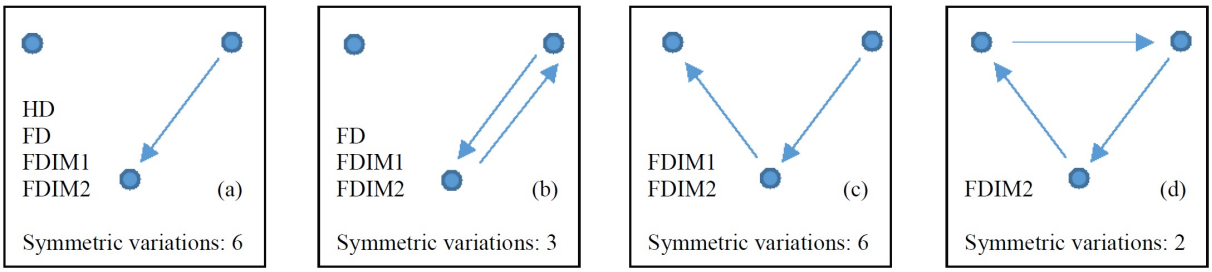


FIGURE 5.3. Allowable (simultaneous) transmissions: (a) single transmission; (b) bidirectional full-duplex transmission; (c) unidirectional full-duplex transmission; (d) an extension of (c) with three nodes in full-duplex simultaneously. Each panel is labelled with the scenarios (HD, FD, FDIM1, FDIM2) under which it is permitted. For each panel, all of the symmetric variations of that pictured are also permitted.

with entries $E_{ij}, 1 \leq i, j \leq 6$. If $E_{ij} = 1$, then links i and j cannot transmit simultaneously, whereas if $E_{ij} = 0$ then links i and j can transmit simultaneously. In this particular example, interference (and hence the collision matrix) is symmetric: that is, if link i collides with link j, then link j also collides with link i.

Scenario 1: Half-Duplex (HD)

See Fig. 5.3(a). In this scenario at most one link can be active at any specific instant (i.e., at most one transmission can take place at any one instant). Every link collides with every other link. Therefore the resulting collision matrix - see Fig. 5.4(a) - is full of 1's, in all off-diagonal positions.

Scenario 2: Full-Duplex (FD)

Scenario 1 is extended to also allow bidirectional full-duplex [52] transmission between two nodes sending and receiving to each other as shown in Figs. 5.2 and 5.3(b). It is supposed that all nodes are equipped with a transmitter and receiver and have self-interference cancellation capability. This is reflected in the collision matrix by setting $E_{12} = E_{21} = 0$, $E_{34} = E_{43} = 0$ and $E_{56} = E_{65} = 0$. Note that in this scenario, all possible combinations of three simultaneous transmissions (i, j, k) result in a collision - see Fig. 5.4(b).

Scenario 3: Full-Duplex with Interference Management - scheme 1 (FDIM1)

Scenario 2 is extended to also allow unidirectional full-duplex [52] transmissions as shown in Figs. 5.2 and 5.3(c). In such a set-up, it is supposed that a given node ($\mathcal{N}_B$ say) may receive from another ($\mathcal{N}_C$ say) and simultaneously transmit to a third node ($\mathcal{N}_A$ say). Of course, the MAC protocol must be extended in some way to allow this possibility. However, the physical layer must also be managed so that $\mathcal{N}_A$ can hear $\mathcal{N}_C$ without interference by the transmission from $\mathcal{N}_B$ to $\mathcal{N}_C$. For example, beamforming / some kind of directional transmission might achieve this capability, and although these are still open research topics, it is supposed that the physical layer can achieve this separation perfectly so that the study can focus on interference effects at the

$$\begin{pmatrix} 0 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 \end{pmatrix} \qquad \begin{pmatrix} 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 \end{pmatrix} \qquad \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}$$

(a) Scenario 1 (HD) (b) Scenario 2 (FD) (c) Scenario 3 & 4 (FDIM1 & FDMI2)

FIGURE 5.4. Link collision matrices $\mathbf{E}$ for each of the scenarios. A '0' in position (i, j) indicates that links i and j may transmit simultaneously and a '1' in position (i, j) indicates that they may not. The collision matrices for the FDIM1 and FDIM2 scenarios are identical.

MAC level. In the collision matrix for this scenario, see Fig. 5.4(c), the only pairs of simultaneous transmissions that are not collision-free are those which either originate or terminate at the same node. (Within the study it is not possible to transmit to two other nodes simultaneously nor to receive from two other nodes simultaneously.)

Scenario 4: Full-Duplex with Interference Management - scheme 2 (FDIM2)

This is a minor extension of Scenario 3 supposing that the MAC protocol is further extended to allow multiple nodes to transmit simultaneously in full-duplex [241] as shown in Fig. 5.3(d). In fact, the collision matrix for this example, see Fig. 5.4(c), is identical to Scenario 3. There are two possible 'complete triangles' of transmissions: e.g., $\mathscr{N}_A \rightarrow \mathscr{N}_B$, $\mathscr{N}_B \rightarrow \mathscr{N}_C$, $\mathscr{N}_C \rightarrow \mathscr{N}_A$, corresponding to $\mathscr{L}_1$, $\mathscr{L}_3$ and $\mathscr{L}_5$. Note $E_{13} = 0$, $E_{35} = 0$ and $E_{51} = 0$. So in simulation, the distinction between FDIM1 and FDIM2 is that FDIM1 has logic to disallow three simultaneous transmissions, even if the collision matrix appears to permit them.

5.2.3 Simulation Methodology

The investigation simulates each of the Scenarios 1-4 using the conceptual simulation model (see Chapter 4). The collision matrix used to model the network was varied to model the respective scenarios. The protocol was adjusted between the half-duplex and full-duplex scenarios with a few additional timesteps added to the *Transmit* state of the conceptual simulation model to reflect the longer carrier sensing overhead and additional management data transmission of the proposed FD-MACs [52, 160, 281]. For the purpose of better illustrating the impact of the technologies discussed, an error prone channel is assumed, here setting $Countdown_{\max} = 14$.

For each Scenario, a large ensemble of simulations are performed, each with a different demand vector (d_A, d_B, d_C) (see Sec. 5.2.1). Specifically, let d_{sat} be the maximum transmission rate of a single half-duplex node transmitting alone in a clear channel. Each of the demands d_A, d_B and d_C is varied independently from 0 to d_{sat} in 40 equal increments, resulting in $41^3 = 68,921$

simulations for each of the four Scenarios. The duration of each simulation is set at $10,000/d_{\text{sat}}$ (i.e., the time needed to transmit $10,000$ half-duplex packets without collisions).

For each individual simulation, statistics on the total number of packets successfully transmitted, the time evolution of queues and the average latency per packet are gathered. These statistics are post-processed with various heuristics (see Sec. 4.3) to decide whether each simulation is within the capacity region, meaning the network and protocol can meet the prescribed demand; or outside the capacity region, meaning that in the large time limit, latency and at least one queue grow without bound.

The results that follow are based on comparisons between Scenarios 1-4 of the proportion of simulated demand vectors that are within the capacity region.

5.2.4 Results

For each of the four scenarios, the simulation results can be presented in the form of a three-dimensional scatter plot, see Fig. 5.5(a), where red markers indicate combinations of demand that are within the capacity region. In this plot $\delta_A := d_A/d_{\text{sat}}$, $\delta_B := d_B/d_{\text{sat}}$ and $\delta_C := d_C/d_{\text{sat}}$ denote non-dimensional demand intensities that range from 0 to 1. The concavity of the capacity region (the volume covered by red dots) is apparent from such plots and represents the loss of efficiency in the channel due to competition between transmitters and the resulting collisions. However, in order to compare the scenarios numerical measures that can be derived from Fig. 5.5(a) are required. Three such measures are used.

Firstly, define total demand intensity $\delta := \delta_A + \delta_B + \delta_C$. Then sections $\delta = \text{const.}$ describe triangular cross sections through Fig. 5.5(a). See Fig. 5.5(b), which helps give a clearer picture of the structure of the capacity region. One may then count (as a function of δ) the proportion of the corresponding triangular area that is within the capacity region, and denote this quantity S. The dependence of S on δ may then be studied and compared across the four scenarios: see Fig. 5.6(a).

Secondly, one may consider the line $\delta_A = \delta_B = \delta_C$ along which the demands are equal. One may then identify the maximum value δ_{cap} (of $\delta_A = \delta_B = \delta_C$) which is within the capacity region, and compare across scenarios. See Fig. 5.6(b).

Finally, one may measure the proportion of the simulations that are within the capacity region - that is, the proportion of the volume (V) $[0,1] \times [0,1] \times [0,1]$ that is within the capacity region - and compare across scenarios. See Fig. 5.6(b).

The point of the various measures is that they allow one to distinguish whether capacity is added by enhancing goodput in symmetric demand situations, or by allowing fresh combinations of highly asymmetric flows.

5.2.5 Analysis

The results of the simulations are summarised in Fig. 5.6. In particular, in Fig. 5.6(a), a higher value of S for a given δ is better in that it means a higher proportion of demand scenarios for

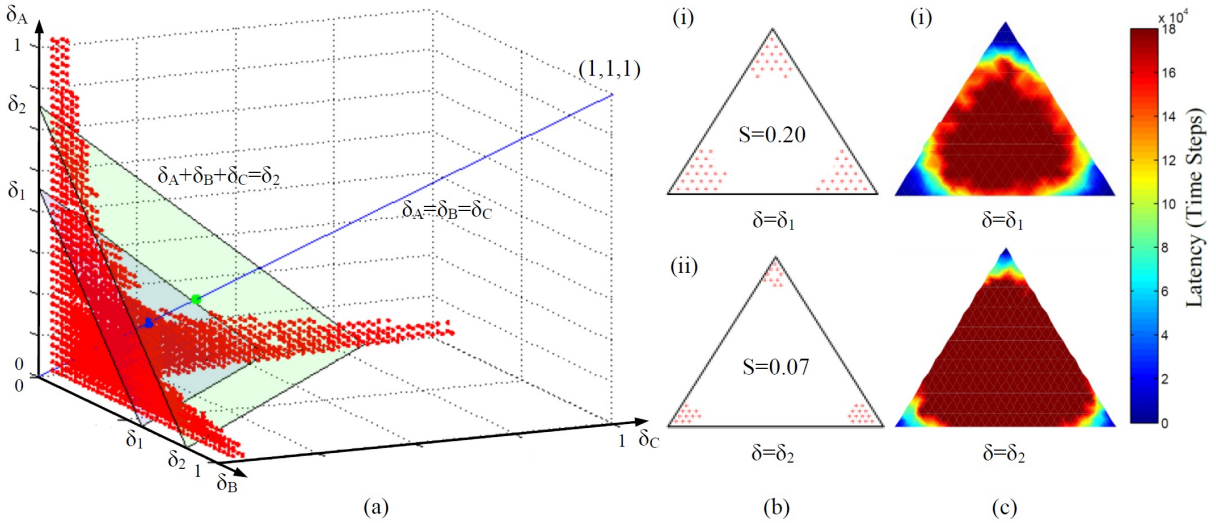
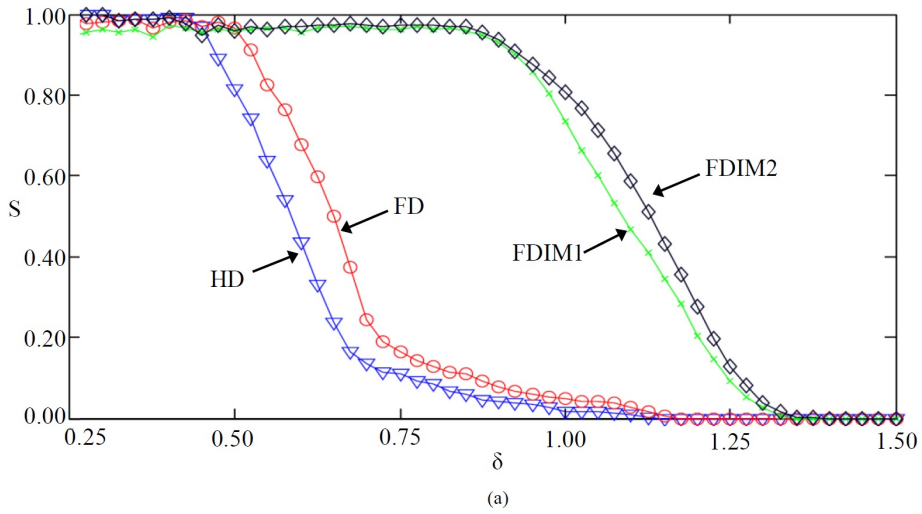


FIGURE 5.5. (a) Capacity region plot for Scenario 1 (HD). Red markers indicate demand combinations within the capacity region. (b) Triangular sections $\delta := \delta_A + \delta_B + \delta_C =$ const. for constants $\delta_1 = 0.67$ and $\delta_2 = 0.82$. S denotes the proportion of the triangle covered with red markers. (c) Corresponding plots of latency, showing how it diverges outside the capacity region.



Scenario	V	δ_{cap}
HD	0.033	0.59
FD	0.045	0.66
FDIM1	0.182	1.17
FDIM2	0.197	1.17

(b)

FIGURE 5.6. Numerical results. (a) Proportion S of triangular sections within the capacity region, as a function of total demand intensity $\delta = \delta_A + \delta_B + \delta_C$, compared across the four Scenarios. (b) Table of the volume V of the capacity region, and its extent δ_{cap} along the line $\delta_A = \delta_B = \delta_C$.

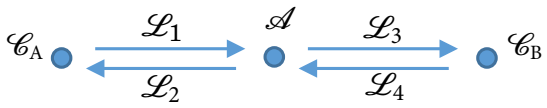


FIGURE 5.7. Simple network setup. Each client $\mathscr{C}_A$ and $\mathscr{C}_B$ is connected to the access point $\mathscr{A}$ by two directional links $\mathscr{L}_1, \mathscr{L}_2, ..., \mathscr{L}_4$.

given total demand δ are within the capacity region. Thus the right-shift of curves as the study progresses through Scenarios 1-4 (from HD to FD to FDIM1 to FDIM2) represents improvements in capacity. However, although each Scenario is an improvement over its predecessor, it is clear that the biggest gain is achieved from FD to FDIM1, that is by the introduction of interference management to full-duplex.

The results, summarised in Fig. 5.6(b), show significant growth in the capacity region's volume V. However, this does not directly relate to the increase in δ_{cap} implying the observed growth in V is in uneven demand combinations. Scenario 2 (FD) was shown to increase V by 37% and δ_{cap} by 11.9% in comparison to Scenario 1 (HD). Supported by the findings of [267], it is anticipated that in a network of more nodes, the gain would be less due to aggregate interference. Implementing Scenario 3 (FDIM1), V grew by 455% and δ_{cap} 98.3% relative to Scenario 1 (HD). Implementing Scenario 4 (FDIM2) showed a further 45% increase in V, a total increase of 500% but the same δ_{cap} increase of 98.3% in comparison to Scenario 1 (HD).

Of course, the application of full-duplex, by allowing two or more simultaneous transmissions, is expected to improve total demand capacity δ_{cap}. It is also expected interference management (Scenarios 3 and 4) should improve matters further since there are larger numbers of combinations of nodes that can transmit simultaneously without collisions. However, the capacity region method shows that the most dramatic improvements in performance are in uneven demand combinations — for example where (say) δ_A and δ_B are high, and δ_C is low. In standard bidirectional full-duplex (Scenario 2) $\mathscr{N}_A$ and $\mathscr{N}_B$ are able to lock into sustained full-duplex transmission with each other, except that the symmetric demand scenario prescribed requires that they each transmit to $\mathscr{N}_C$ as well. With the addition of interference management (Scenarios 3 and 4), $\mathscr{N}_A$ (say) can break off to transmit to $\mathscr{N}_C$, yet $\mathscr{N}_B$ can continue to transmit to $\mathscr{N}_A$.

5.3 Study 2: A Performance Evaluation of the Impact of Full-Duplex Capable Nodes on Common Problems in Wireless Networks

5.3.1 Simple Network Setup

This study again considers a simple three-node network. In this instance, two of those nodes are client devices that can communicate directly with the third node, a wireless access point,

via a single hop, see Fig. 5.7. The client devices cannot communicate with each other directly. This topology is used to represent two common issues in current communications networks; bottlenecks and hidden nodes. When a WLAN network is saturated, a bottleneck occurs at the access point. This occurs because the access point is subject to the same protocol, and thus has the same probability of gaining access to the channel for transmission as each client while serving more demands. The hidden node problem (see Chapter 3 p.58) occurs when a half-duplex receiver cannot notify other nodes in the network that it is currently receiving a signal and thus other nodes attempt to transmit to the receiver. Implementing full-duplex has the potential to mitigate against these issues. The study that follows demonstrates the potential performance gains full-duplex nodes can offer.

The study considers no routing, each client ($\mathscr{C}_A, \mathscr{C}_B$) can transmit to the access point ($\mathscr{A}$) directly and hence there are four possible transmissions (links) that are denoted $\mathscr{L}_1, \mathscr{L}_2, ..., \mathscr{L}_4$. It thus follows one may prescribe demand (number of packets per unit time) on the network in terms of requested flows $d_1, d_2, ..., d_4$ for each of the four links respectively. However, to simplify the computational analysis, the study is restricted to the special symmetric case $d_1 =: d_{\mathscr{C}_A}, d_2 = d_3 =: d_{\mathscr{A}}$, and $d_4 =: d_{\mathscr{C}_B}$, and hence need only scan through the three-dimensional demand space $(d_{\mathscr{C}_A}, d_{\mathscr{A}}, d_{\mathscr{C}_B})$.

Thus it is assumed each node (including the access point) has a single queue of packets awaiting transmission with Poisson arrival rate: $d_{\mathscr{C}_A}, d_{\mathscr{A}}, d_{\mathscr{C}_B}$ respectively. The queue on $\mathscr{A}$ will consist on average of an equal mix of packets bound for $\mathscr{C}_A$ and $\mathscr{C}_B$. Further, it is assumed that there is no active queue management. Hence, regardless of the destination, packets queued at the access point are served in order of arrival.

A further description of the network modelling process is explained in Sec. 5.3.2 and assumptions concerning packet collisions etc for the bottleneck and hidden node problems are explained in Sec. 5.3.3.

At various stages of the modelling process (as with the previous study) two forms of full-duplex transmission are considered: unidirectional, where the access point can receive from one client and simultaneously transmit to another, and bidirectional, where the access point can receive from and transmit to the same client simultaneously [52]. These are shown in Fig. 5.8, and how they are modelled is explained in Sec. 5.3.2.

5.3.2 Operational Scenarios

The study will now consider collisions between the various links within the network and how they might be mitigated by full-duplex techniques. The bottleneck and hidden node problems will be considered with three operational scenarios in relation to each of them: half-duplex access point and half-duplex clients (HDAP-HDC), full-duplex access point and half-duplex clients (FDAP-HDC), and full-duplex access point and full-duplex clients (FDAP-FDC). In fact, these are the only scenarios of significance. For example, it would not be of use to consider a half-duplex

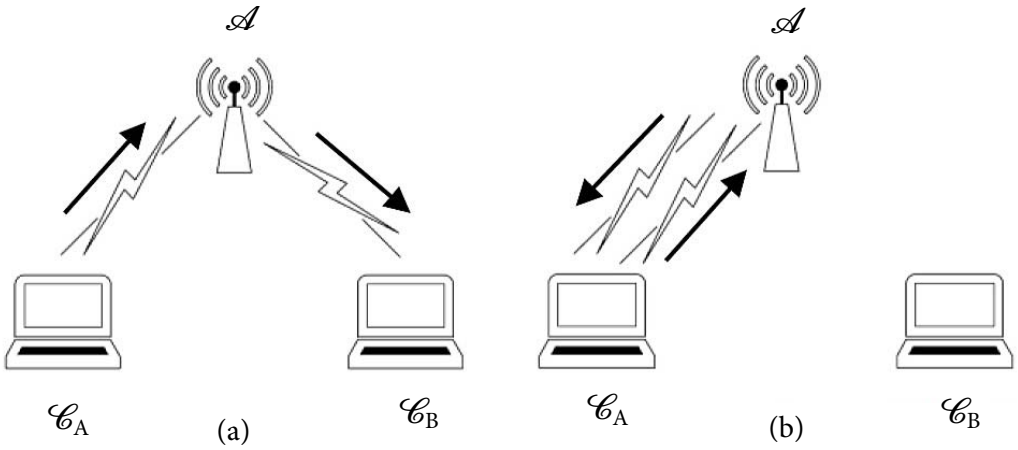


FIGURE 5.8. Simultaneous transmissions occurring in full-duplex (a) Unidirectional full-duplex (b) Bidirectional full-duplex.

access point and full-duplex clients as the clients would be forced to communicate with the access point via a half-duplex protocol, making this scenario the same as HDAP-HDC.

Throughout a simple deterministic binary interference model is supposed, so that if two incompatible links transmit simultaneously, there will be a collision with probability one. Thus here a four-by-four collision matrix **E** is used (see Fig. 5.9) to describe each operational scenario, with entries $E_{ij}, 1 \leq i, j \leq 4$. If $E_{ij} = 1$, then link i cannot transmit simultaneously with j, whereas if $E_{ij} = 0$ then link i can transmit simultaneously with j. In this particular example, interference (and hence the collision matrix) is symmetric: that is, if link i collides with link j, then link j also collides with link i, i.e., $E_{ij} = E_{ji}$.

Scenario 1: Half-Duplex Access Point and Half-Duplex Clients (HDAP-HDC)

In this scenario at most one link can be active at any specific instant (i.e., at most one transmission can take place at any one instant). Every link collides with every other link. Therefore the resulting collision matrix - see Fig. 5.9(a) - is full of 1's, in all off-diagonal positions.

Scenario 2: Full-Duplex Access Point and Half-Duplex Clients (FDAP-HDC)

Suppose that Scenario 1 is modified such that the access point is now full-duplex capable. This, therefore, allows unidirectional full-duplex [52] transmissions as shown in Fig. 5.8(a). It is supposed that the access point is equipped with a transmitter and receiver that have self-interference cancellation capability. In such a set-up, the access point may receive from a client ($\mathscr{C}_A$ say) and simultaneously transmit to the other client ($\mathscr{C}_B$). Of course, the MAC protocol must be extended in some way to allow this possibility. With the introduction of a full-duplex access point to this scenario, certain pairs of simultaneous transmissions are now allowed to occur collision-free. This is reflected in the collision matrix by setting $E_{13} = E_{24} = 0$, and $E_{31} = E_{42} = 0$, see Fig. 5.9(b). (It is supposed not possible for the access point to transmit to two clients simultaneously nor to receive from two clients simultaneously.)

$$\begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix} \qquad \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix} \qquad \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

(a) (b) (c)

FIGURE 5.9. Link collision matrix **E**. a) HDAP-HDC b) FDAP-HDC c) FDAP-FDC.

Scenario 3: Full-Duplex Access Point and Full-Duplex Clients (FDAP-FDC)

It is supposed that Scenario 2 is further modified such that the clients, as well as the access point, are now full-duplex capable. This therefore additionally allows bidirectional full-duplex [52] transmissions as shown in Fig. 5.8(b). It is supposed that the clients and access point are equipped with a transmitter and receiver and have self-interference cancellation capability. In such a set-up, the access point ($\mathscr{A}$) may receive from a client ($\mathscr{C}_A$ say) and simultaneously transmit to the same client ($\mathscr{C}_A$). In the collision matrix for this scenario, see Fig. 5.9(c), further pairs of simultaneous transmissions are allowed to occur collision-free. This is implemented in the collision matrix by setting $E_{12} = E_{34} = 0$, and $E_{21} = E_{43} = 0$.

5.3.3 Modelling the Bottleneck and Hidden Node Problems

To investigate the impact of full-duplex nodes on the bottleneck and hidden node problems, simulations based on the three Scenarios described in Sec. 5.3.2 are performed. The conceptual simulation model described in Chapter 4 is again applied.

Scenario 1 represents the bottleneck problem in its simplest form. The clients are both serving just one demand to the access point while the access point is serving demand to each of the two clients. Thus a bottleneck will occur at the access point. Scenarios 2 and 3 are modifications of Scenario 1 introducing full-duplex nodes. For this part, it is assumed for each Scenario that the nodes have a global knowledge of the network, i.e., no hidden nodes.

The premise of the hidden node problem stems from the issue that links do not have a global knowledge of the network and therefore are not necessarily aware of transmissions on other links with which they may collide. This can be modelled by introducing a second matrix that this thesis calls the 'knowledge of the network' matrix **F** (see Fig. 5.10).

This is a four-by-four matrix to describe each link's awareness of the network around it, with entries $F_{ij}, 1 \leq i, j \leq 4$. If $F_{ij} = 1$, then link i is aware of j and will not transmit simultaneously, whereas if $F_{ij} = 0$ then link i either can transmit simultaneously or is unaware of j (cannot hear its transmissions). As with the collision matrices in this particular example, we suppose that the network topology (and hence the matrix **F**) is symmetric: that is, if link i has knowledge of link j, then link j also has knowledge of link i i.e., $F_{ij} = F_{ji}$. A lack of knowledge of each other's

$$\begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix} \qquad \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \qquad \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

(a) (b) (c)

FIGURE 5.10. Knowledge of the network link matrix - **F**. a) HDAP-HDC b) FDAP-HDC c) FDAP-FDC.

transmissions can result in the clients transmitting simultaneously, each not realising the other is transmitting, causing a collision.

Consider the topology depicted in Fig. 5.7: the access point is able to hear either of the clients transmitting, similarly both of the clients are able to hear the access point transmitting. The hidden node problem occurs as a result of $\mathscr{C}_A$ not being able to hear transmissions by $\mathscr{C}_B$ and vice versa. This is captured in matrices **F**, which are otherwise the same as the three collision matrices, by setting $F_{14} = 0$ and $F_{41} = 0$ (see Fig. 5.10).

In summary: the hidden node problem is modelled by implementing the three described Scenarios. The study assumes for each Scenario that the nodes have knowledge of the network via **F** but are subject to collisions via **E**. For the three simulations **F** is modelled using the 'knowledge of the network' matrix in Figs. 5.10 (a), (b), and (c) respectively, however the collision matrices **E** are as in Figs. 5.9 (a), (b), and (c) respectively.

5.3.4 Simulation Methodology

The investigation simulates each of the Scenarios 1-3 for both the bottleneck problem and the hidden node problem using the conceptual simulation model. By this, it is meant that the collision matrix and knowledge of the network link matrix used to model the network were varied for the respective problem and Scenario being investigated (Secs. 5.3.2 and 5.3.3) and the protocol was adjusted between the half-duplex and full-duplex appropriately (as in the previous study, $Countdown_{\max} = 14$).

For each Scenario, a large ensemble of simulations was performed, each with a different demand vector $(d_{\mathscr{C}_A}, d_{\mathscr{A}}, d_{\mathscr{C}_B})$. As with the previous study, let d_{sat} be the maximum transmission rate of a single half-duplex client or access point transmitting alone in a clear channel. Each of the demands $d_{\mathscr{C}_A}$, $d_{\mathscr{A}}$ and $d_{\mathscr{C}_B}$ is varied independently from 0 to d_{sat} in 40 equal increments, resulting in $41^3 = 68{,}921$ simulations for each of the two problem's three Scenarios. The duration of each simulation is set at $10{,}000/d_{\text{sat}}$ (i.e., the time needed to transmit 10,000 half-duplex packets without collisions).

For each individual simulation, statistics on the total number of packets successfully transmitted, the error rate, the time evolution of queues, and the average latency per packet are gathered.

The results that follow are based on comparisons between Scenarios 1-3 for the bottleneck problem and the hidden node problem.

5.3.5 Results

For each of the two problems investigated, the three scenarios' simulation results can be presented in the form of a three-dimensional scatter plot, see Fig. 5.11(a), where red markers indicate combinations of demand that are within the capacity region. In this plot $\delta_{\mathscr{C}_A} := d_{\mathscr{C}_A}/d_{\text{sat}}$, $\delta_{\mathscr{A}} := d_{\mathscr{A}}/d_{\text{sat}}$ and $\delta_{\mathscr{C}_B} := d_{\mathscr{C}_B}/d_{\text{sat}}$ denote non-dimensional demand intensities that range from 0 to 1. The concavity of the capacity region (the volume covered by red dots) is apparent from such plots, and represents the loss of efficiency in the channel due to competition between transmitters and the resulting collisions. However, numerical measures that can be derived from Fig. 5.11(a) are required in order to compare the scenarios.

For the purpose of analysis, firstly consider the line $\delta_{\mathscr{C}_A} = \delta_{\mathscr{A}} = \delta_{\mathscr{C}_B}$ along which the demands are equal. Secondly, define aggregate demand intensity $\delta := \delta_{\mathscr{C}_A} + \delta_{\mathscr{A}} + \delta_{\mathscr{C}_B}$. Then δ = const. defines triangular cross sections through Fig. 5.11(a).

One may examine the ratio of successful transmissions to error rate (on the line $\delta_{\mathscr{C}_A} = \delta_{\mathscr{A}} = \delta_{\mathscr{C}_B}$) and compare across the two problems and three scenarios. See Fig. 5.11(b). From this figure the demand (δ) at which each client and the access point reaches saturation is apparent.

One may also examine latency across the triangular sections δ_1 and δ_2, and compare across the three scenarios for each of the two problems. See Fig. 5.11(c).

Fig. 5.12(a) helps give a clearer picture of the structure of the capacity region. One may then count (as a function of δ) the proportion of the corresponding triangular area that is within the capacity region, and denote this quantity S. The dependence of S on δ may then be studied and compared across the four scenarios: see Fig. 5.12(b).

Further, the proportion of the simulations that are within the capacity region is measured - that is, the proportion of the volume (V) $[0,1] \times [0,1] \times [0,1]$ that is within the capacity region: see Fig. 5.12(c).

These various measures allow one to distinguish whether capacity is added by enhancing goodput in symmetric demand situations, or by allowing fresh combinations of highly asymmetric flows.

5.3.6 Analysis

The findings of this study are summarised in Figs. 5.11 and 5.12.

Fig. 5.11(c) demonstrates the impact of full-duplex for demand intensity $\delta = 0.65$ and $\delta = 0.80$. The figure indicates how the latency around nodes with full-duplex capability is significantly reduced from half-duplex nodes. The hidden node plots show as expected lower latency around the access point (i.e., for combinations of demands in which the access point has the greatest share), which is able to hear all nodes in the network and therefore is not subject to the problem of hidden

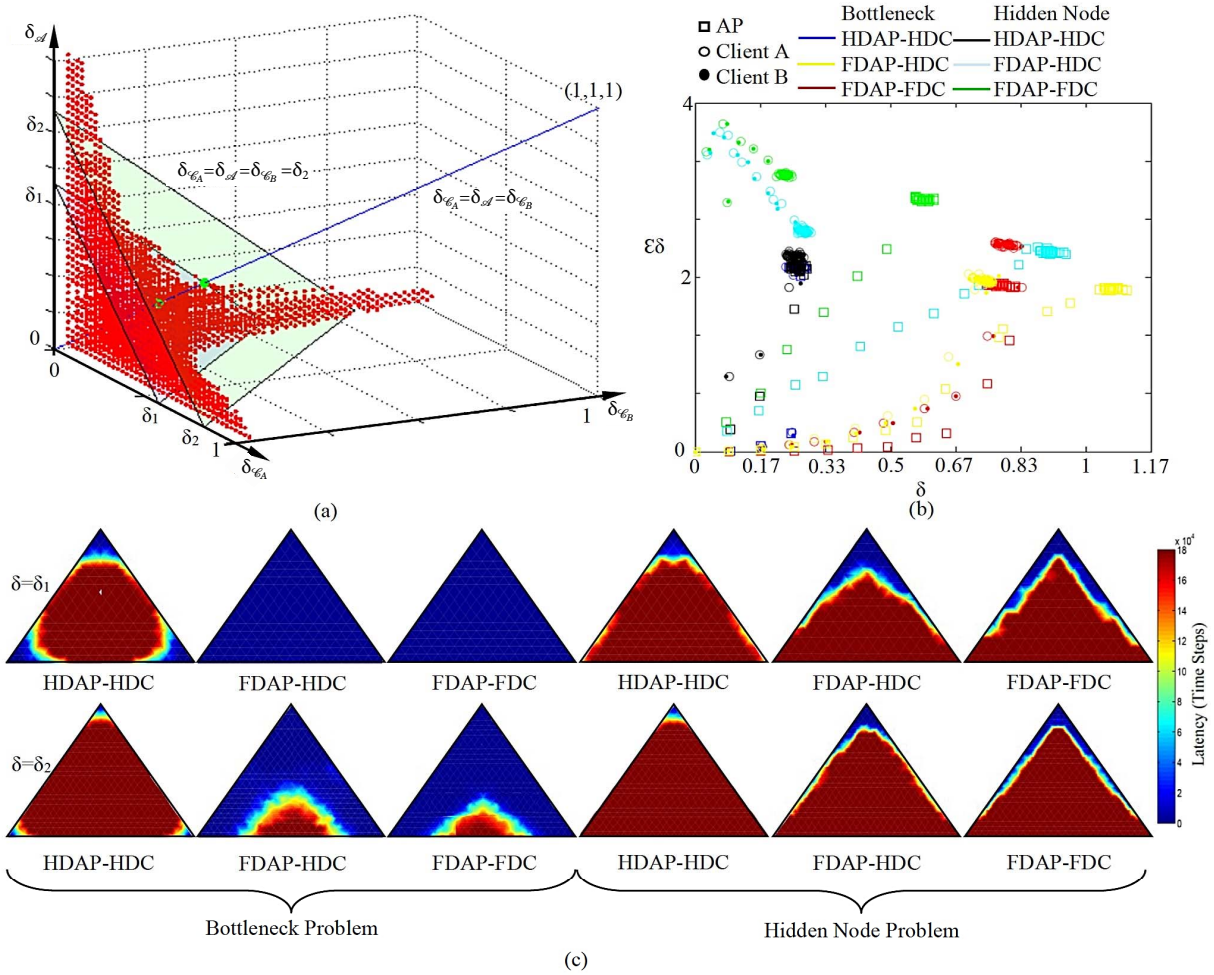
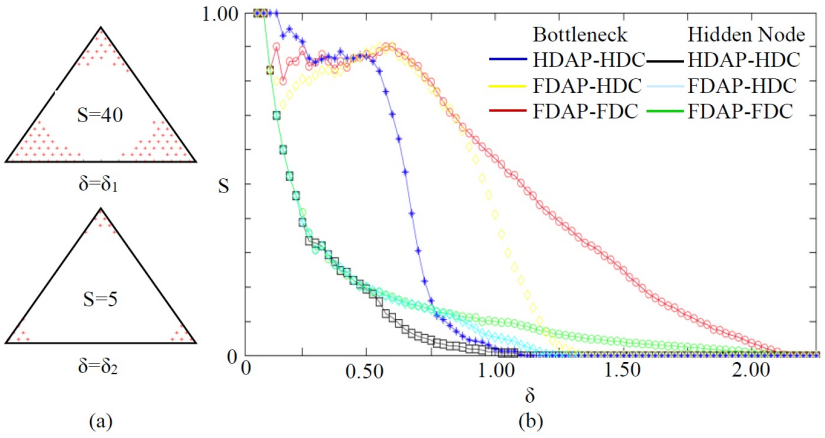


FIGURE 5.11. (a) Capacity region plot for Bottleneck Problem HDAP-HDC. Red markers indicate demand combinations within the capacity region. (b) A plot of demand (δ) against ($\varepsilon\delta$), where ε is the error rate i.e., the number of attempted transmissions per successful transmission, along the line $\delta_{\mathscr{C}_A} = \delta_{\mathscr{A}} = \delta_{\mathscr{C}_B}$. (c) Triangular sections $\delta := \delta_{\mathscr{C}_A} + \delta_{\mathscr{A}} + \delta_{\mathscr{C}_B} = \text{const.}$ for constant $\delta_1 = 0.65$ (top row) and $\delta_2 = 0.80$ (bottom row). Plots show latency across the corresponding section for the two problems and three Scenarios investigated. The top of each triangle represents the access point $\mathscr{A}$ and the left and right corners clients $\mathscr{C}_A$ and $\mathscr{C}_B$ respectively.



Problem	Scenario	V
Hidden Node	HDAP-HDC	0.01
Hidden Node	FDAP-HDC	0.02
Hidden Node	FDAP-FDC	0.05
Bottleneck	HDAP-HDC	0.04
Bottleneck	FDAP-HDC	0.12
Bottleneck	FDAP-FDC	0.27

(c)

FIGURE 5.12. (a) Triangular sections $\delta := \delta_{\mathscr{C}_A} + \delta_{\mathscr{A}} + \delta_{\mathscr{C}_B} = \text{const.}$ (see Fig. 5.11) for constant $\delta_1 = 0.65$ and $\delta_2 = 0.80$ for Bottleneck HDAP-HDC simulation. S denotes the proportion of the triangle covered with red markers. (b) Proportions S of triangular section within the capacity region, as a function of total demand intensity $\delta = \delta_{\mathscr{C}_A} + \delta_{\mathscr{A}} + \delta_{\mathscr{C}_B}$, compared across three scenarios for each of the two problems. (c) Table of volume V of the capacity region.

nodes as the clients are. Comparing the hidden node plots for $\delta = \delta_1$ Scenario 2 (FDAP-HDC) seems to show a larger area of lower latency around the access point (the top corner of the triangle) in comparison to Scenario 3 (FDAP-FDC). However, Scenario 3 shows the area of low latency extending with a clearer definition from the access point to the clients down the links of the triangle. The same is true, but less visibly clear, for $\delta = \delta_2$. This suggests a lower goodput from the access point in Scenario 3 than Scenario 2 but improved goodput from the clients. Comparing the bottleneck problem simulations, the Scenario 1 (HDAP-HDC) plot shows a symmetrical plot with latency equal on all corners of the triangle. This reflects the nodes' awareness of other in the network and equal probability of accessing the channel. Introducing full-duplex shows a significant improvement in performance, with the most significant improvement occurring around the access point.

Consider Fig. 5.11(b), plotting demand against error rate along the line $\delta_{\mathscr{C}_A} = \delta_{\mathscr{A}} = \delta_{\mathscr{C}_B}$. For the bottleneck problem simulations, the access points and the clients follow a broadly similar pattern with the error rate increases as the traffic is increased. Scenarios 2 (FDAP-HDC) and 3 (FDAP-FDC) show an increase in the value of δ at which the network reaches saturation compared to Scenario 1 (HDAP-HDC). In Scenario 1 both the access point and clients follow the same trend and reach saturation at the same value of δ. For Scenario 2 the goodput at which the access point reaches saturation is higher than the clients with roughly similar error rates. For Scenario 3 the client and the access point reach saturation at approximately the same value of δ, however, the error rate from clients' transmissions is higher. The access point is saturated in

Scenario 2 at higher δ than Scenario 3 whereas the clients for both Scenarios saturate at similar values of δ.

Continue to analyse Fig. 5.11(b). For the hidden node problem at low δ, the error rate ε of clients, transmissions in Scenarios 2 (FDAP-HDC) and 3 (FDAP-FDC) is shown to be high, significantly higher than those in Scenario 1 (HDAP-HDC). As δ increases for Scenarios 2 and 3, ε for clients decreases and ε for the access point increases. Further, in Scenarios where the traffic at the access point is saturated, or close to saturated, the goodput is shown to be significantly increased and the ratio of successful to failed transmissions is improved using a full-duplex access point relative to a half-duplex access point. It is shown that in Scenario 3 there is a higher error rate (failed packets per successful transmission) than in Scenario 2. Further, the saturation of Scenario 2 occurs at a higher value of δ for the clients and a significantly higher value of δ for the access point (more than three times) than Scenario 3.

Consider Fig. 5.12(b): a higher value of S for a given δ is better in that it means a higher proportion of demand scenarios for a given total demand δ are within the capacity region. Therefore a right shift of the curves indicates an improvement in capacity. Here for both the bottleneck and hidden node problems, an increase can be observed from Scenario 1 (HDAP-HDC) to Scenario 2 (FDAP-HDC) and a further increase to Scenario 3 (FDAP-FDC). This is summarised in Fig. 5.12(c), which shows the capacity regions' volumes V. This result differs from the findings above considering the line $\delta_{\mathscr{C}_A} = \delta_{\mathscr{A}} = \delta_{\mathscr{C}_B}$ where Scenario 2 performed better than Scenario 3. This indicates that much of the growth in V is in uneven demand combinations — for example where (say) $\delta_{\mathscr{A}}$ and $\delta_{\mathscr{C}_A}$ are high, and $\delta_{\mathscr{C}_B}$ is low.

5.4 Discussion

The results of the first study (Sec. 5.2) are summarised in Fig. 5.6. The results illustrate that the introduction of full-duplex will provide an increase in capacity from half-duplex. However, the gain is only slight. A much more significant gain is achieved from full-duplex to full-duplex combined with interference management. This example assumes all nodes in the network have a knowledge of each other, i.e., there are no hidden nodes.

Consider the hidden node problem: for all scenarios simulated the knowledge of the network link matrices differs from the collision matrix which indicates a risk that certain incompatible links may attempt to transmit simultaneously (i.e., both clients attempt to transmit simultaneously to the access point) regardless of the applied protocol. Due to the ability of a full-duplex access point to transmit whilst receiving, it can dictate to the clients which other links can transmit simultaneously in full-duplex. Further, the busy signal from the access point silences other clients that may potentially cause a collision with the transmitting client. As a result, the problem of hidden nodes is, to an extent, mitigated by full-duplex on the access point.

It is clear from Fig. 5.12 that regardless of the protocol or the number of full-duplex capable

nodes, the presence of hidden nodes significantly decreases the network capacity. For both the bottleneck and hidden node simulations, the introduction of full-duplex nodes shows an increase in the capacity region, thus contributing (although in some Scenarios only slightly) to addressing the two problems. As the findings of [267] show, it is anticipated that in a network with a greater number of clients, the gain achieved from full-duplex would be less due to additional aggregate interference.

Fig. 5.11(b) shows that in hidden node Scenarios, as traffic is increased, the error rate of the clients decreases (up to the point of saturation). It is apparent from the results that in a situation where the demand at the access point is low, the goodput of the network could be improved by the access point simply transmitting noise (i.e., a busy tone) to silence other clients.

In the second study, both the bottleneck and hidden node problem simulations have shown Scenario 2 (FDAP-HDC) to reach saturation at a higher value of δ than Scenario 3 (FDAP-FDC), and have a lower error rate at saturation when considering equal demand to all nodes. Further, for both problems, the clients of Scenario 2 and 3 have reached saturation at similar values of δ, however, in both cases, the error rate for Scenario 2 has been lower. This result is analogous with observations made by, among others, Bianchi [29] in his analysis of 802.11 networks. It is commonly known that several random access schemes exhibit an unstable behaviour. As the load on the network increases, the goodput rises up to a maximum value. Further increasing the load on the network can lead eventually to a decrease in system performance. In comparison to Scenario 2, for an equivalent value of δ, the number of attempted transmissions from the clients in Scenario 3 (all clients full-duplex) is significantly higher. However, this increase in the number of transmission attempts is found to have a negative effect on the overall network performance, resulting in an increased number of collisions, reduced performance of the access point and lower goodput.

The results show that at saturated (or sufficiently high) traffic the introduction of full-duplex nodes can significantly reduce the effect of hidden nodes. The improvement in performance from all nodes will allow for increased goodput in networks with full-duplex access points. Further, the greater improved performance of the access points over the clients with the introduction of full-duplex will enable networks to better cope with bottlenecks formerly caused by one access point serving multiple clients.

Of course, the results presented in both examples could be refined by increasing the number of demand combinations (i.e., use more than 40 increments from zero to d_{sat}) and sampling error (apparent as irregularity / noise in Figs. 5.5(C) and 5.11(C)) could be reduced by increasing the run time for each individual simulation. However, both such refinements would significantly increase the total CPU time required for this study. In any case, the aggregate measures V and δ_{cap} are relatively stable to misclassification of individual simulations, and so the present studies seem safe in their overall conclusions concerning the capacity gains of FD-MAC.

Further, normative testing of the conclusions could be carried out using more detailed sim-

ulation models (e.g., with NS3 [196]) of the MAC protocol and the physical layer. However, the computational cost of such normative simulations is very high and will result in impractical run times if a full scan of demand space (in the manner of the presented study) were to be attempted.

5.5 Conclusion

This chapter contributes to knowledge in the following ways:

CTK 5.1 Demonstrating the potential capacity gain that full-duplex can provide applied in various forms to a highly simplified mesh network set-up.

The first study in this chapter shows that full-duplex alone can increase the capacity of a network, however, when full-duplex is combined with appropriate interference management, the gain is much more significant. Existing work speculates at this result but this study is the first to quantify it.

CTK 5.2 Identifying that the majority of the capacity gain achieved by full-duplex occurs at asymmetric demand combinations.

This would not have been identified by just considering equal or saturated demands.

CTK 5.3 Showing that introducing full-duplex access points alone mitigates against the problem of bottlenecks, reduces the impact of hidden nodes and can increase the capacity of a network. It identifies that when full-duplex access points are able to work with full-duplex clients, the capacity gain is much more significant, however, it is shown that much of this capacity gain occurs at uneven demand combinations.

CTK 5.4 Identifying limitations of full-duplex. When the demand to all nodes is equally high, the introduction of full-duplex capability to clients is shown to increase the number of transmission attempts resulting in a significantly increased number of collisions and reduced network performance.

CTK 5.5 Discovering that at low traffic levels, a full-duplex access point may improve goodput by simply transmitting a busy tone to silence other transmissions whilst it receives, mitigating against the hidden node problem.

This chapter conducted studies on two topologies. The first study investigated the capacity gain available from the application of full-duplex to wireless mesh networks. The study considered the simplest possible meaningful example: a three-node mesh network, in which all nodes are within range of each other.

Four different operational scenarios were modelled: one half-duplex and three full-duplex with differing levels of interference management. Simulation results have shown that the application of full-duplex alone can improve the network capacity however the increase is much more significant

when combined with interference management techniques. The capacity region statistics that were used enable a finer understanding of potential network performance gains than mere measurement of saturated flows. Much of the capacity gain is shown to occur at uneven demand combinations.

A second study investigated the performance gain available from the introduction of full-duplex to a simple network with one access point and two clients. The study considered two recognised problems in communications networks: bottlenecks and hidden nodes. The two problems for three different operational scenarios are modelled: one with two half-duplex clients and a half-duplex access point, one with two half-duplex clients and a full-duplex access point and one with two full-duplex clients and a full-duplex access point.

Simulation results have shown that the introduction of full-duplex access points alone can improve the network capacity, reduce bottlenecks and the undesirable effect of hidden nodes. The addition of full-duplex clients further improves the capacity of the overall network, however much of this gain is at uneven demand combinations. For even demand combinations, the addition of full-duplex clients increased the number of attempted transmissions, however, also increased the error rate such that the overall network performance decreased. Again, statistics provided in relation to the capacity region enable improved understanding of potential network performance gains compared to the simple measurement of saturated flows. To mitigate against hidden nodes at low traffic levels, a full-duplex access point may improve goodput by simply transmitting a busy tone to silence other transmissions whilst it receives.

Further work should consider how the conclusions of these studies extend to larger and more realistic network examples with more nodes. A further complication is that an exhaustive search of the (very high dimensional) demand space will not be computationally tractable, and more refined search techniques will need to be adopted.

CHAPTER 6

A Pragmatic Approach to Clear Channel Assessment Threshold Adaptation and Transmission Power Control For Performance Gain in CSMA/CA WLANs

This chapter is based on: W. Jones, R. E. Wilson, A. Doufexi, and M. Sooriyabandara "A Pragmatic Approach to Clear Channel Assessment Threshold Adaptation and Transmission Power Control For Performance Gain in CSMA/CA WLANs" IEEE Transactions on Mobile Computing, Jan. 2019.

This chapter proposes a practical set of rules for adapting CCA and TP parameters by simulating ensembles of randomly generated wireless networks and collecting throughput statistics. The rules' performances depend strongly on network topology, with increases in throughput (and goodput) in many cases. However, networks with a high clustering coefficient are often adversely affected by the adaptations. But simulations of small-scale networks show that apparently adverse adaptations may still yield benefits for uneven demand combinations. Finally, and surprisingly, it is found that throughput is not usually correlated with the number of hidden or exposed nodes in any non-trivial network set-up.

6.1 Introduction

The densification of CSMA/CA WLANs has helped improve coverage, however, it has also increased interference and led to poor spatial reuse [282]. One idea, to avoid simultaneous transmissions in neighbouring cells being suppressed unnecessarily, is to adapt the CCA threshold.

The opportunity is best explained by the simple example of Fig. 6.1 consisting of four nodes, two clients ($\mathscr{C}_{1,2}$) and two access points ($\mathscr{A}_{1,2}$). Cell 1 ($\mathscr{C}_1$ and $\mathscr{A}_1$) is adjacent to Cell 2 ($\mathscr{C}_2$ and $\mathscr{A}_2$). $\mathscr{C}_1$ and $\mathscr{C}_2$ sense each others' transmissions and will not transmit simultaneously even

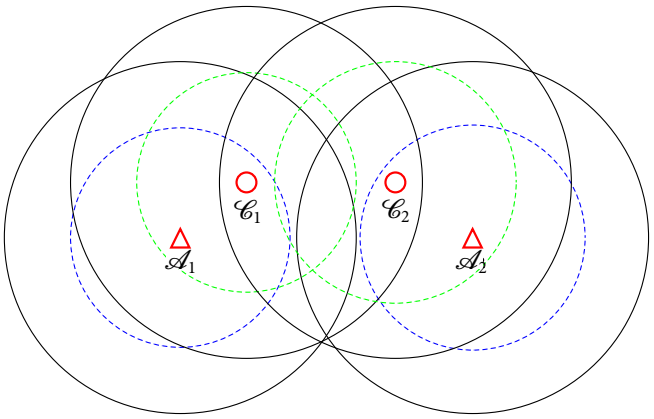


FIGURE 6.1. Basic CCA threshold and TP adaptation. Large black solid circles surrounding each node represent legacy CCA threshold or TP. Blue and green dashed circles represent reduced CCA sensitivity/TP for APs and clients respectively. The diagram illustrates the opportunity to prevent suppression of transmissions in adjacent cells by reducing CCA sensitivity or TP.

though their signals as received at the APs would not interfere with each other (exposed nodes [127]). If the clients were to adjust their CCA threshold sensitivity from the legacy [149] (i.e., fixed original value before any adaptation of the parameter may take place — black solid circles in Fig. 6.1) to a higher threshold (less sensitive value, green dashed circles in Fig. 6.1) they would be deaf to each others transmissions and thus could transmit simultaneously.

Alternatively, parallel transmissions can be achieved by adapting the TP of the clients (maintaining legacy CCA threshold). Again considering Fig. 6.1, if $\mathscr{C}_1$ and $\mathscr{C}_2$ modify their TP from the legacy (now represented by the black solid circles) to a reduced value (the green dashed circles) the transmissions are too quiet for the clients to sense each other but the APs can still receive them.

An extensive survey by Thorpe [248] summarises the CCA literature. Similarly, Chiang's book [56] provides a comprehensive review of TP control research. Much of the work discussed in [248] and [56] focuses on finding the best possible solution for one particular fixed network topology. Uncoordinated optimisation of the CCA and TP parameters across a high number of nodes is unlikely to ever reach a true system optimum throughput: with all nodes aiming to maximise their own throughput, a select number will dominate the channel at the expense of others. Rather, this chapter defines and investigates a set of heuristic rules for adapting CCA threshold and TP and evaluates their efficiency by collecting throughput statistics.

This chapter is organised around a simulation study in two parts as follows. The first proposes a set of rules to adapt the CCA threshold, the TP or a combination of the two, and studies an ensemble of randomly generated networks at legacy CCA and TP compared with each of the rules applied, at saturated traffic conditions (Sec. 6.2). The results are analysed (Sec. 6.2.1) and

show that adapting TP alone typically results in the most significant throughput gain and often improves the fairness [125] of the network. Further, it is observed that networks with a low clustering coefficient are most likely to achieve a throughput gain (Sec. 6.2.2). It is shown that the average throughput gain, for some rules, is higher when only applied to networks with a low clustering coefficient [22, 268]. To better understand these observations and the impact of the proposed rules, two simple networks of five nodes are studied (Sec. 6.3) applying similar CCA and TP rules as proposed in Sec. 6.2. Due to the reduced number of nodes it is possible to explore a larger parameter space: instead of focusing on equal saturated demands, the study investigates vectors of demand that can be met by the system (Sec. 6.3.2), and presents results in terms of the capacity region [151] (Sec. 6.3.3). This metric allows the findings to be analysed (Sec. 6.3.4) identifying gains in performance at asymmetric demand combinations. Further, it provides an improved understanding of the relationship between throughput performance and fairness, and how the rules impact these.

The key outputs of this chapter (discussed in Sec. 6.4) are an evaluation of the performance of the proposed rules, and an improved understanding of the capability of CCA and TP adaptation, in terms of their ability to reduce cross cell interference, improve spatial reuse and increase system throughput. This work highlights that adapting CCA threshold or TP can impact nodes' ability to carrier sense each other, thus improving spatial reuse. By adapting TP, simultaneous transmissions are enabled which would previously not have been possible due to interference at receivers. In networks with large numbers of nodes, uncoordinated adaptation of the CCA and TP parameters is unlikely to reach a true system optimal throughput. However, it is shown that simple heuristic rules can bring some performance gains. How the approach may be implemented in dynamic networks is discussed.

6.2 Randomly Generated Networks

Consider a physical set-up where the nodes (clients and access points) are positioned within a square subset $[0,L]\times[0,L]$ of the (x,y) plane. Firstly, clients $\mathscr{C}_i$, with $i=1,2,\ldots,n$, are uniformly randomly distributed across the square, which is divided into cells organised around four access points $\mathscr{A}_j$, $j=1,2,3,4$, positioned at $(1/2,1/2)\pm(1/4,0)\pm(0,1/4)+\mathbf{e}_j$, where the $\mathbf{e}_j$ are independently drawn random variables from the uniform distribution on $[-L/10,+L/10]\times[-L/10,+L/10]$ (see Fig. 6.2). The idea is that the APs are (roughly) symmetrically positioned, but the clients might be spread quite unevenly.

Clients i and APs j are then allocated default values for their CCA and TP. The simplified approach followed (see [73, 142]) is to model CCA sensitivity and TP as separate distance thresholds, rather than as a single power density threshold experienced by the receiver. Specifically, it is supposed that a node $\mathscr{N}_A$ (client or AP) is assumed able to transmit to, or interfere with,

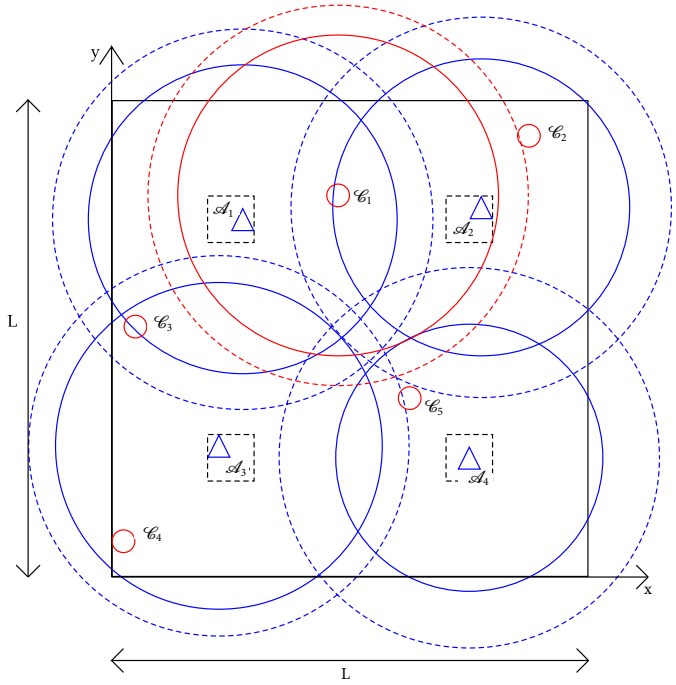


FIGURE 6.2. Example of a randomly generated network, the subset $[0,L]\times[0,L]$ of the (x,y) horizontal plane (large black box). The rectangle is divided into four cells organised around $\mathscr{A}_{1,2,3,4}$ (blue triangles) at $[L/4,L/4]\pm\mathbf{e}$, $[3L/4,L/4]\pm\mathbf{e}$, $[L/4,3L/4]\pm\mathbf{e}$, $[3L/4,3L/4]\pm\mathbf{e}$, where $\mathbf{e}$ is in each case an independently drawn random variable from the uniform distribution on $[-L/10,+L/10]\times[-L/10,+L/10]$ (small dashed boxes). $\mathscr{C}_{1,2,\ldots,5}$ (small red circles) are randomly distributed across the rectangle and four cells. Each client and AP is at the centre of a circle representing its CCA threshold and TP. To preserve clarity the legacy CCA and TP of the APs (blue dashed and blue solid circles respectively centring at each AP) and the legacy CCA and TP of just $\mathscr{C}_1$ (large red dashed and large red solid circles respectively centring at $\mathscr{C}_1$) are shown. It is clear $\mathscr{C}_1$ could be heard by both $\mathscr{A}_1$ and $\mathscr{A}_2$ at legacy conditions.

another node (client or AP) $\mathscr{N}_B$ if

$$\|\mathbf{x}_{\mathscr{N}_A}-\mathbf{x}_{\mathscr{N}_B}\|<\mathrm{TP}_{\mathscr{N}_A},\ \mathrm{CCA}_{\mathscr{N}_B} \tag{6.1}$$

i.e., the receiver must be within 'transmission range' of the transmitter, and the transmitter must be within an 'audible range' of the receiver. In fact, (6.1) simplifies the true state of affairs, which in these terms is

$$\|\mathbf{x}_{\mathscr{N}_A}-\mathbf{x}_{\mathscr{N}_B}\|^2<\mathrm{TP}_{\mathscr{N}_A}\mathrm{CCA}_{\mathscr{N}_B}. \tag{6.2}$$

This is because power density at the receiver scales with transmit power $\mathrm{TP}_{\mathscr{N}_A}$ and the inverse square of the separation $\|\mathbf{x}_{\mathscr{N}_A}-\mathbf{x}_{\mathscr{N}_B}\|$ of the transmitter and receiver. Note further that CCA

sensitivity measured as a distance should scale reciprocally with CCA measured as a power density threshold. The difficulty is that (6.1) and (6.2) are not equivalent in situations where both $\mathrm{TP}_{\mathscr{N}_A}$ and $\mathrm{CCA}_{\mathscr{N}_B}$ are varied, because (6.1) implies (6.2) but not vice versa. However, the approximate formulation (6.1) enables an intuitive graphical understanding of range that (6.1) does not, that proves especially useful in Sec. 6.3. In addition, having tested the accuracy of this simplification in terms of the similarity of the binary collision matrices they generate for the systems, it was found that over 99% of the matrices agree. The simplified approach is thus to proceed with (6.1) as an exact model.

For simplicity, experiments consider a fixed legacy CCA sensitivity range of $4L/10$ for all nodes and suppose the legacy TP value is drawn from the uniform distribution on $[3L/10, 4L/10]$.

Client $\mathscr{C}_i$ then associates to the closest access point $\mathscr{A}_j$ with

$$j = \arg\min_k \|\mathbf{x}_{\mathscr{A}_k} - \mathbf{x}_{\mathscr{C}_i}\| \tag{6.3}$$

such that

$$\|\mathbf{x}_{\mathscr{A}_k} - \mathbf{x}_{\mathscr{C}_i}\| < \mathrm{TP}_{\mathscr{C}_i},\ \mathrm{CCA}_{\mathscr{C}_i},\ \mathrm{TP}_{\mathscr{A}_k},\ \mathrm{CCA}_{\mathscr{A}_k}, \tag{6.4}$$

so that the access point and client can hear each others' transmissions bi-directionally. In the case where there is no access point in range, the client is randomly re-positioned and the association considered again. A link is formed between $\mathscr{C}_i$ and $\mathscr{A}_j$ along which transmissions occur.

The study area dimensions and CCA/TP values are selected to model a four cell network, where cell borders overlap and hence are likely to be subject to the detrimental effects of cross-cell interference. For example, by considering the example network of five clients shown in Fig. 6.2, it is clear that $\mathscr{C}_1$ will be audible to both $\mathscr{A}_1$ and $\mathscr{A}_2$ at legacy CCA and TP.

If two clients attempt to transmit to the same AP simultaneously, a collision will occur and one or both transmissions will fail. Similarly, if the received interference from a transmitter in a neighbouring cell is too great at an AP, a transmission may fail. Further, if $\mathscr{C}_1$ (say) could hear $\mathscr{C}_2$ (say) transmitting, $\mathscr{C}_1$ would wait for $\mathscr{C}_2$ to finish before transmitting itself. Note $\mathscr{C}_1$ could hear $\mathscr{C}_2$ if $\|\mathbf{x}_{\mathscr{C}_1} - \mathbf{x}_{\mathscr{C}_2}\| < \mathrm{TP}_{\mathscr{C}_2}, \mathrm{CCA}_{\mathscr{C}_1}$. Transmitters do not have a global knowledge of the network and therefore are not necessarily aware of transmissions on other links with which they may collide (hidden node problem — see Fig. 3.6 p.59) or transmissions which suppress them from transmitting, with which they could successfully transmit in parallel (exposed nodes — see Fig. 3.7 p.60). Thus, a collision matrix $\mathbf{E}$ and 'knowledge of the network' matrix $\mathbf{F}$ (both n by n) are used to describe each operational scenario (see Secs. 5.2.2 and 5.3.2). In the examples, interference (and hence the collision matrix) is not always symmetric: that is, if link i collides with link j, in some cases only one link transmission may fail, i.e., $E_{ij} = 1$ but $E_{ji} = 0$. Similarly, each nodes awareness of the network topology (and hence the matrix $\mathbf{F}$) is not symmetric: that is, if link i has knowledge of link j, then link j does not necessarily have knowledge of link i i.e., $F_{ij} = 1$ but $F_{ji} = 0$.

This chapter proposes five different adaptation rules as follows which are applied in turn:

R1 All nodes of all cells have legacy CCA threshold value and legacy TP.

R2 All nodes of all cells reduce their CCA sensitivity (increasing threshold) to the minimum value at which they can still carrier sense all nodes of their cell that they were able to at legacy CCA value. A 5% margin is added. Legacy TP for all nodes.

R3 All nodes of all cells reduce their TP to the minimum value at which they can still transmit to, and be carrier sensed by, all nodes of their cell they were able to at legacy TP. A 5% margin is added. Legacy CCA for all nodes.

R4 All nodes of all cells reduce their CCA sensitivity by half the potential difference to the minimum value at which they can still carrier sense all nodes of their cell which they were able to at legacy CCA value. Simultaneously, all nodes of all cells reduce their TP by half the potential difference to the minimum value at which they can still transmit to, and be carrier sensed by, all nodes of their cell which they were able to at legacy TP. A 5% margin is added to both parameters.

R5 At random, cells choose to reduce their CCA sensitivity (as R2) or reduce their TP (as R3).

The investigation generates $10,000$ networks of n clients (and hence n links) with legacy TP and CCA threshold values (i.e., R1). Ten network densities of $n = 5, 10, ..., 50$, are considered and therefore $100,000$ networks each defining a pair of matrices $\mathbf{E}$ and $\mathbf{F}$. Further developing each of the legacy networks, the described CCA threshold and TP adaptations (R2-R5) are applied in turn generating four further $\mathbf{E}$ and $\mathbf{F}$ matrices pairs to capture them. In total this produces $500,000$ pairs of matrices $\mathbf{E}$ and $\mathbf{F}$. Using the conceptual simulation model described in Chapter 4 (here $Countdown_{\max} = 2$), data traffic is simulated for each of these networks (i.e., $500,000$ individual simulations).

The CSMA/CA MAC protocol [29] requires completion of a two-way handshake in order for one transmission to be successful. This study is interested in understanding the extent of hidden node [127] issues when multiple clients unable to carrier sense each other attempt to transmit to the same AP. Similarly, the study seeks to gain insight into the impact of exposed nodes [127] when clients in neighbouring cells suppress each other's transmissions. In the particular topological setup of this chapter where AP positions are restricted, down-link transmissions are less prone to these issues. Hence the study considers only up-link demand, that is APs do not transmit data to clients, just clarification that the client is clear to send (CTS) and acknowledgement (ACK) of received packets. The general findings are however applicable to both up and down-link. It is supposed that there is no routing and that each packet has a journey that consists of a single hop. It follows that one may prescribe demand (number of packets per unit time) on the network in terms of requested flows d_1, d_2, ..., d_n for each link respectively. Thus, each client is assumed to have a single queue of packets awaiting transmission, with Poisson arrivals. For each simulation, demand to all transmitters is set at d_{sat}, i.e., the maximum transmission rate of a

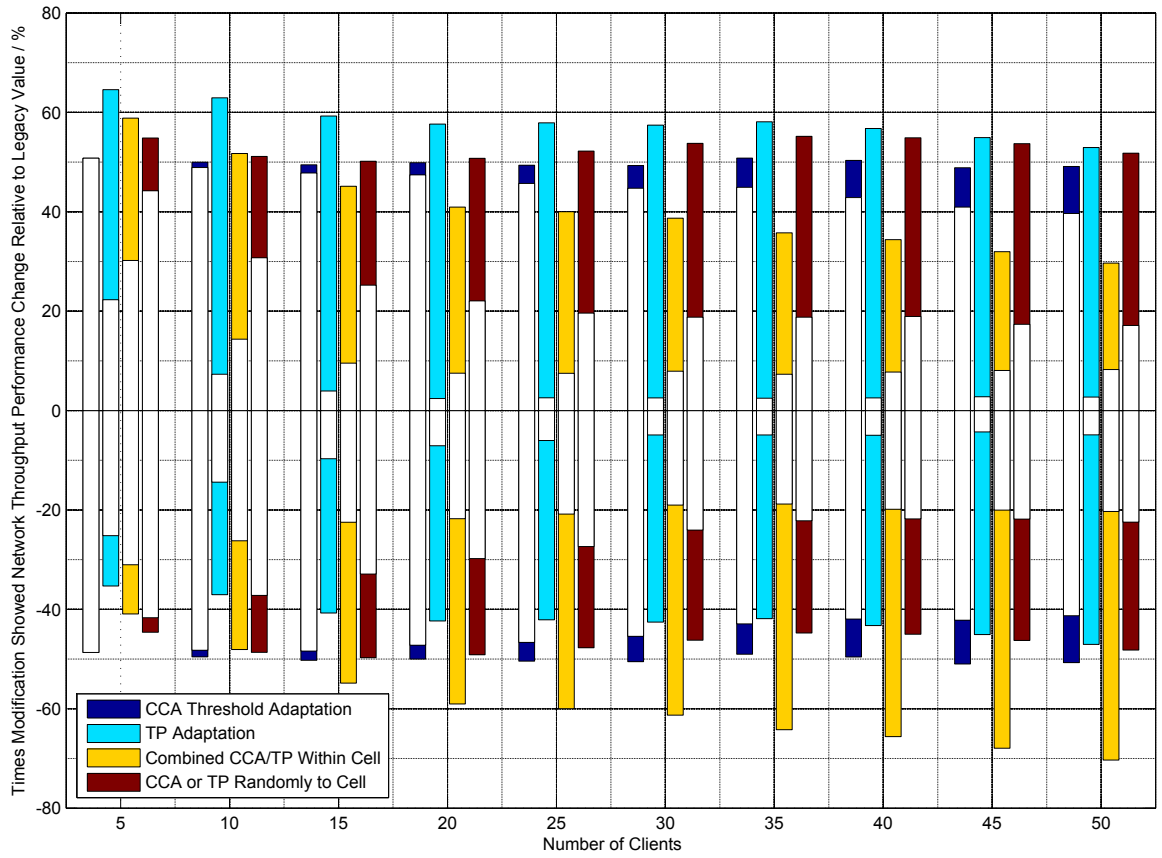


FIGURE 6.3. Performance impact of applied modifications. Bars indicate the number of occasions (%) that the modification led to an increase or decrease in throughput. The coloured sections of the bars indicate a change greater than 10%.

single transmitter transmitting alone in a clear channel. The duration of each simulation is set at $100,000/d_{\text{sat}}$ (i.e., the time needed to transmit 100,000 packets without collisions).

Each network and the five rules under which that network is simulated are compared relative to each other. The total number of packets successfully transmitted and the fairness (based on Jain's fairness index [125]) under each of those five rules for the same (saturated) traffic demand and time period is measured. The results that follow (Sec. 6.2.1) are based on comparisons between rules R1-R5 applied in turn to each of the 10,000 networks for n clients.

6.2.1 Randomly Generated Networks Results and Analysis

For each network, the impact on throughput of each adaptation relative to the legacy condition is explored (i.e., R2-R5 relative to R1 described earlier). Fig. 6.3 illustrates this, plotting four bars for each n clients indicating the percentage of adaptations to the 10,000 legacy networks simulated that resulted in an increased or decreased throughput change. The coloured section of

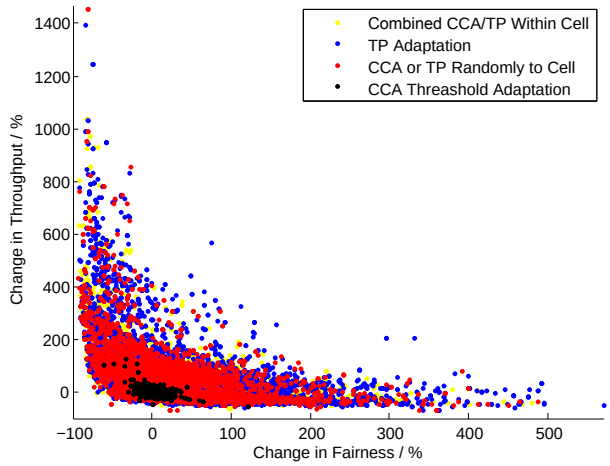


FIGURE 6.4. Impact of adaptation on throughput and fairness relative to legacy for 30 clients.

the bar indicates throughput change greater than 10%.

Consider Fig. 6.3: CCA threshold adaptation (as R2) is shown to have an approximately even probability of generating an increase or decrease in throughput and only a small number of the increases or decreases achieve greater than 10% change. Adapting TP (as R3) showed a higher probability of a throughput gain than loss in all cases, most significantly at five and ten clients. A correlation for the proportion of TP results showing an improvement in throughput decreasing with increasing network density can clearly be seen. A similar correlation can be seen when combining CCA and TP as (R4). This combination only realised a higher probability of a throughput gain for five and ten clients, whereas greater numbers of clients saw a higher probability of a loss. Applying CCA and TP adaptation randomly to different cells (as R5) caused an increase in throughput in more cases than it caused a decrease relative to the legacy for all network densities tested, however, the improvement was only marginal. The most significant improvements occurred at five clients and between thirty to forty-five clients.

Fig 6.4 plots the change in fairness [125] and throughput for each of the adaptations (R2-R5) relative to the 10,000 legacy (R1) networks of thirty clients. The example is typical for any number of clients. When the applied adaptations produce a throughput performance improvement, fairness often decreased, and when fairness increased the network throughput often decreased. The implication is that the adaptation is helping some clients achieve a better throughput, potentially at the expense of others.

The range of impact of each particular adaptation relative to the legacy is visible in Fig. 6.4. The CCA threshold adaptation results are concentrated, showing the smallest range of throughput and fairness change. The result of CCA and TP adaptation randomly applied to different cells show an increased range from CCA alone, but both the TP and CCA/TP combined adaptations show much greater variation, with the potential for significant gains in throughput ($\approx$ 1400%

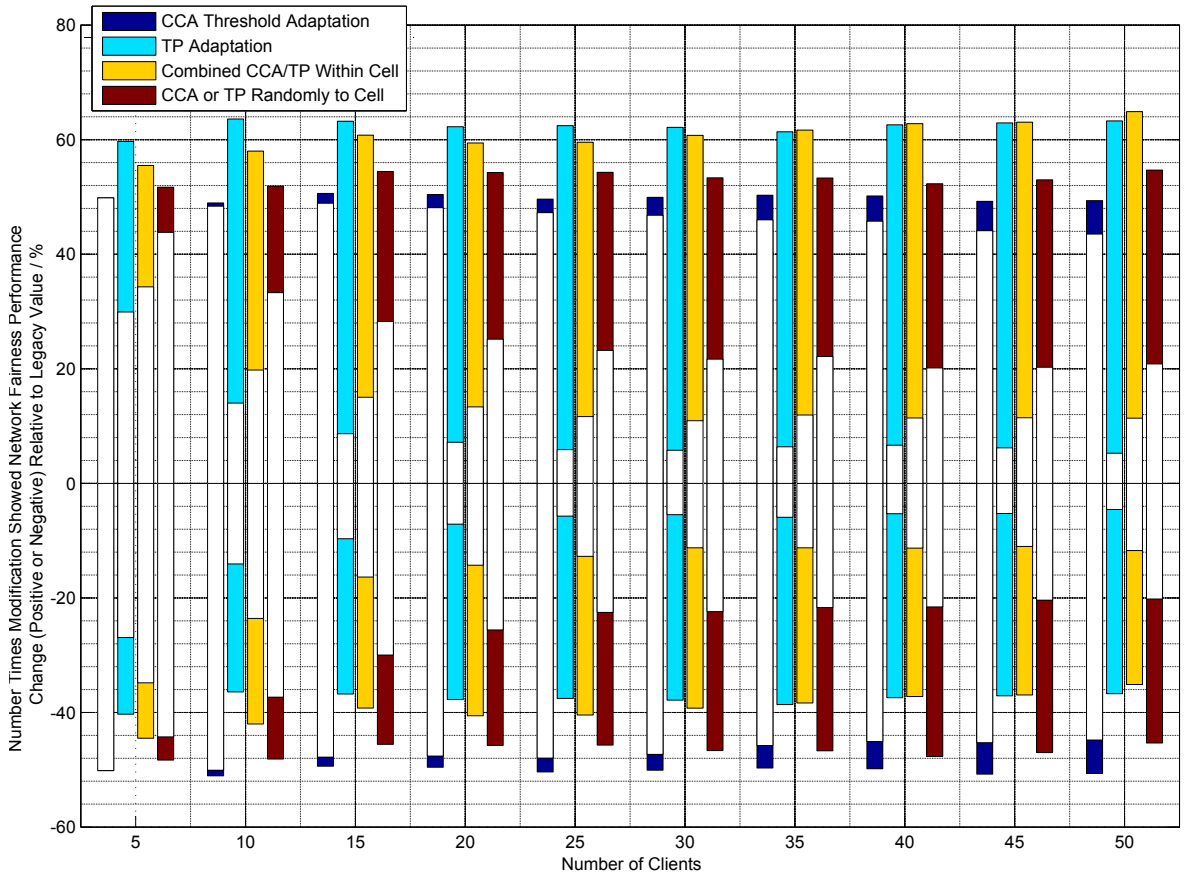


FIGURE 6.5. Fairness impact of applied modifications. Bars indicate the number of occasions (%) that the modification led to an increase or decrease in fairness. The coloured sections of the bars indicate a change greater than 10%.

increase is the most extreme example). Fig. 6.4 reflects the indications of Fig. 6.3 which shows CCA adaptation leading to an increase in throughput on an almost equal number of occasions as it led to a decrease. The figure shows TP adaptation to achieve the biggest change in throughput and the same can be shown for all numbers of clients.

Fig. 6.5 illustrates for each number of clients the percentage of adaptations to the 10,000 legacy networks simulated that resulted in an increased or decreased fairness change (R2-R5 relative to R1). The coloured section of the bar indicates a change greater than 10%. As with throughput, CCA adaptation alone was shown to have an approximately equal probability of achieving a fairness gain or loss. All other adaptations typically led to a fairness improvement on more occasions than a loss, with TP achieving improved fairness most frequently for up to thirty-five clients. For forty and forty-five clients TP and combined CCA/TP achieve a fairness improvement on an approximately equal number of occasions. At fifty clients combined CCA/TP most frequently achieved a fairness improvement.

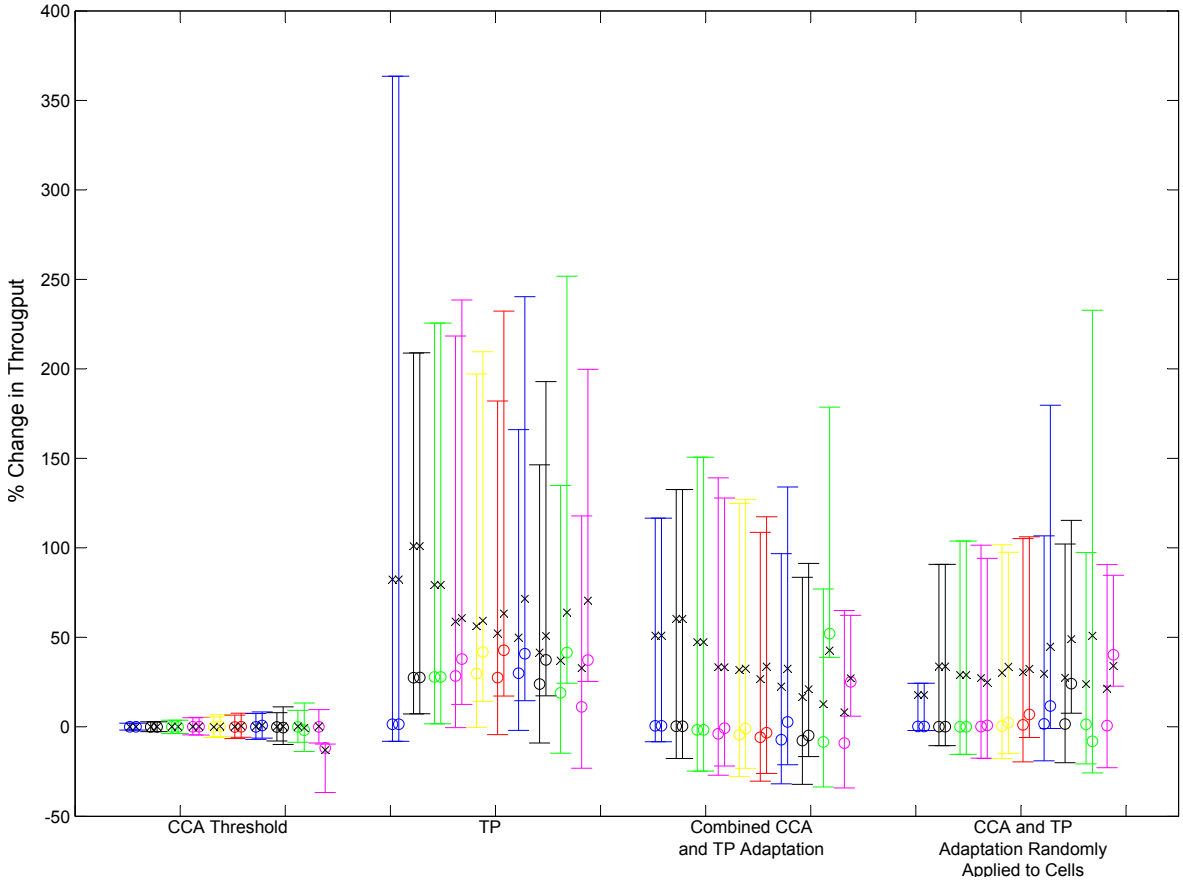


FIGURE 6.6. Performance improvement / reduction range for the 80_{th} percentile, for all adaptation types. For each adaptation type, the group of plots shows 10 coloured pairs of plots for each number of clients, $n = 5, 10, ..., 50$ from left to right. For each range, the circle shows the median value, the black cross shows the mean. The left plot of each pair shows the initial result. The right plot of each pair shows the revised range, mean and median when the adaptation is applied only to networks with clustering coefficient below a threshold value (see Sec. 6.2.2). The potential improvement using the clustering coefficient is visible for 20 to 50 clients.

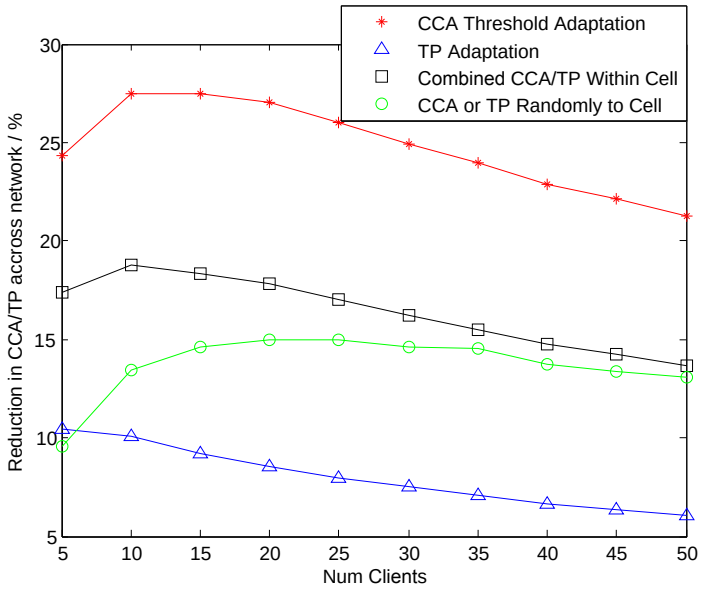
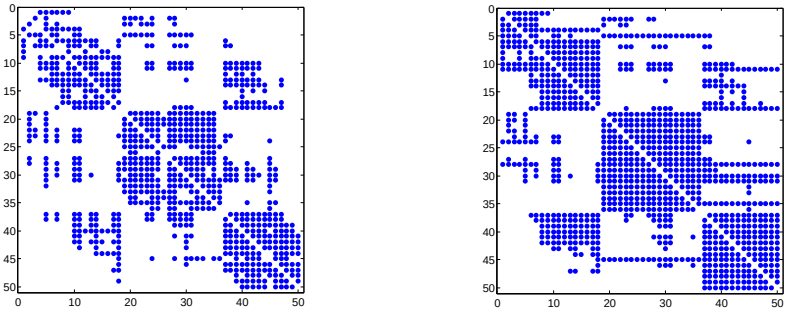


FIGURE 6.7. Mean power reduction for all networks for all adaptation types, normalised for the number of nodes. The plot shows the combined percentage by which both circles surrounding each node defining CCA sensitivity and TP are reduced in the model when each adaptation is applied.

A small number of simulations show very extreme changes from the legacy when an adaptation is applied (see Fig. 6.4). Fig. 6.6 shows the 80th percentile range of throughput changes as a result of the applied adaptations and indicates the mean and median. CCA adaptation shows a small range ($\approx 15\%$) of variations equally distributed positively and negatively either side of zero performance change, reflecting the indications of Fig. 6.3. The TP adaptation shows the potentially highest increase in throughput in Fig. 6.6 for all numbers of clients. The two methodologies that combine CCA and TP adaptation have a similar range of impacts. However, Fig. 6.3 illustrates how differently they perform with combined CCA/TP in each cell achieving a throughput loss on more occasions than that of randomly applying CCA or TP adaptation to different cells. TP is shown to be the most beneficial adaptation for all networks with the most significant throughput improvement at lower density.

Fig. 6.7 shows the total reduction in CCA and TP as a result of the applied adaptations. This figure should be considered in parallel with Figs. 6.3, 6.5 and 6.6. The percentage reduction equates to the cumulative amount by which the radius of the two circles surrounding each node in the model, defining their CCA sensitivity and TP, are reduced from the legacy values. The average change shows a decrease for each adaptation, tending to zero with increasing client density. The figure illustrates the limited change these adaptations are able to achieve as network density is increased, however Figs. 6.3 and 6.6 indicate that this small adaptation can make a significant impact.



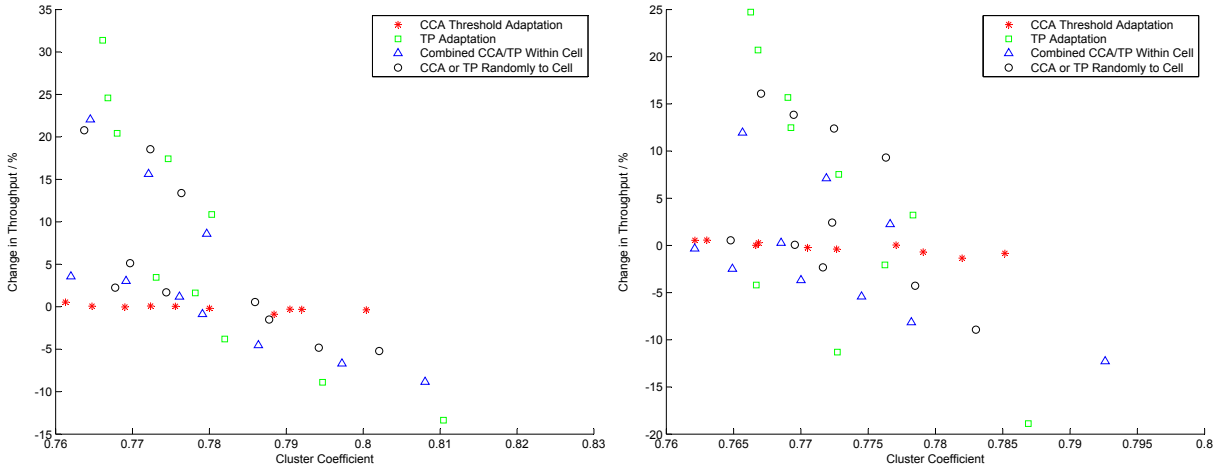
(a) Knowledge Matrix - **F** (b) Collision Matrix - **E**

FIGURE 6.8. An example plot of **F** and **E** matrices for 50 link topology. Blue dots represent a one, no dot represents a zero.

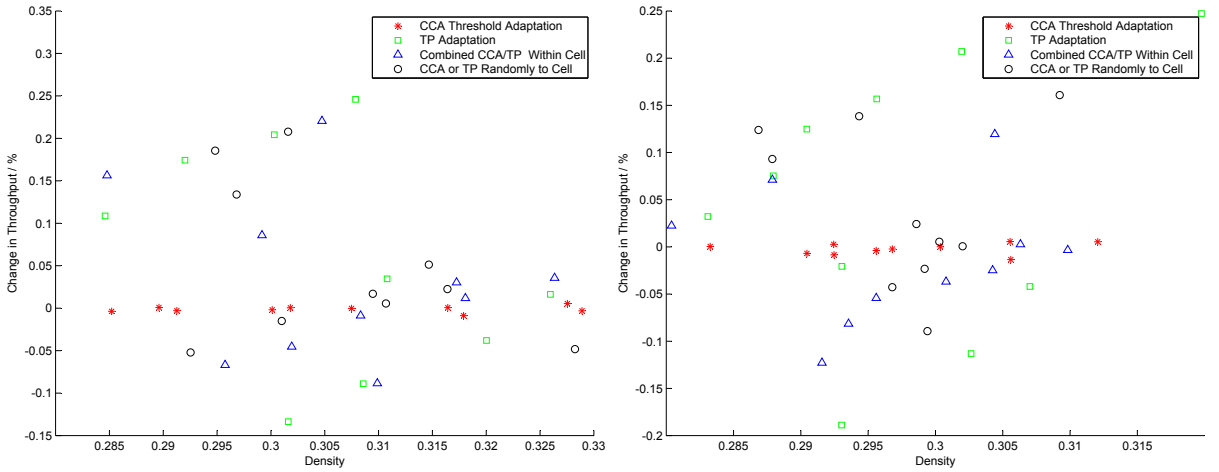
6.2.2 Analysis of Matrix Characteristics

The matrices (**E** and **F**) underlying each simulation, for all numbers of clients, were analysed to identify network characteristics contributing to the performance impact of the applied adaptations (R2-R5 relative to R1 in Sec. 6.2). The matrices (represented in Fig. 6.8) were ordered by their associated AP. The four square clusters visible along the diagonal represent the clients associated with the four APs. The size of each cluster indicates the number of clients associated with that AP. Marked points away from the diagonal indicate clients on the border of two cells. Asymmetry, number of hidden nodes, number of suppressed transmitters, clustering coefficient [22, 268] and other measures were considered.

For all **E** matrices, prior to any adaptation being applied (legacy conditions (R1)), the mean local clustering coefficient was calculated [22, 268]. For each of the four adaptations (R2-R5) the throughput change achieved after the adaptation was ordered from highest to lowest. These were then divided into ten equal-sized groups i.e., the top 10% of performance improvements to the lowest 10% in intervals of 10%. For each of these ten groups the mean cluster coefficient (of the legacy networks) and mean throughput performance improvement were plotted against each other (see Figs. 6.9(a) and 6.9(b)). A correlation shows the percentage change in throughput decreases with increasing clustering coefficient. It is possible to approximately identify a clustering coefficient value for each adaptation and network density above which the application of each adaptation (R2-R5) has a negative effect on throughput from the legacy (R1). This same trend can be identified by comparison of throughput change with **E** or **F**. For reference the relationship between the density of matrix **E**, i.e., proportion of the collision matrix filled with 1's, against throughput change is shown (Figs. 6.9(c) and 6.9(d)) to demonstrate there is no clear trend, confirming that the clustering coefficient is not by itself a consequence in the collision matrix density. The same correlation, for twenty-five and forty-five clients, can be shown similarly in the other size networks investigated. Similar investigations were conducted comparing the number of hidden or exposed nodes in each network with the throughput improvement, however, no trend



(a) Clustering Coefficient vs throughput Change 25 Clients (b) Clustering Coefficient vs throughput Change 45 Clients



(c) Density vs throughput Change 25 Clients (d) Density vs throughput Change 45 Clients

FIGURE 6.9. The mean clustering coefficient and density of **E** matrices compared to mean throughput performance improvement for the top 10% of performing networks to the lowest 10% of performing networks in intervals of 10%.

could be identified.

The identified relationship between clustering coefficient and throughput change (Fig. 6.9) enables an enhanced benefit from TP and CCA adaptations. The set of $100,000$ randomly generated legacy networks (R1) (generated as described in Sec. 6.2) are re-analysed only implementing the adaptation (R2-R5) for a particular number of clients if the cluster coefficient is below a defined threshold, estimated from Figs. 6.9(a) and 6.9(b) (and equivalent figures for other numbers of clients) as the clustering coefficient where the plotted (assumed linear correlation) change in throughput values approximately equals zero. Above that threshold, legacy conditions are maintained and therefore the number of simulations to which the adaptation was applied was reduced. Of those to which the adaptation was applied, for twenty to fifty clients an improvement in accuracy was observed (i.e., the proportion of times the implementation led to an improvement in throughput relative to the legacy when implementing was improved from that shown in Fig. 6.3). No improvement could be shown for five, ten or fifteen clients. For twenty to fifty clients, the mean performance change of TP and the two combined CCA and TP adaptations were improved, most significantly at higher numbers of clients. The revised performance change range, mean and median are added to Fig. 6.6, and the improvement is clearly visible with the number of wrong implementations reduced and the average performance changes improved.

6.3 Simple Topology Networks

To better understand the effects observed in some of the large-scale randomly generated networks, the impact of CCA threshold and TP adaptation are further investigated on two highly simplified networks (see Figs. 6.10 and 6.11). The two networks are similar to those considered in [282] consisting of two APs and three Clients. The only difference between the two is as follows. In Fig. 6.11, $\mathscr{C}_1$ is positioned slightly further away from $\mathscr{A}_1$ and $\mathscr{C}_2$ than in Fig. 6.10. This makes no difference to the clients association with APs, but the change impacts on $\mathscr{C}_1$ and $\mathscr{A}_1$'s ability to adapt their CCA threshold or TP. For both networks (Figs. 6.10 and 6.11), as with the larger network examples, the application of CCA threshold and TP adaptation and a combination of the two are investigated. The study that follows highlights the topological constraints of the interference management techniques and helps further understanding of the results presented in the previous sections.

Some basic assumptions are made about both networks:

- $\mathscr{C}_1$ and $\mathscr{C}_2$ are associated with $\mathscr{A}_1$. These three nodes make up Cell 1. $\mathscr{C}_3$ is associated with $\mathscr{A}_2$. These two nodes make up Cell 2.
- As for the random network examples, only up-link demand from Clients to APs is considered. APs do not transmit data to clients, just CTS and ACK. Thus, the network topologies can be captured by a series of three links; connecting $\mathscr{C}_1$ to $\mathscr{A}_1$, $\mathscr{C}_2$ to $\mathscr{A}_1$ and $\mathscr{C}_3$ to $\mathscr{A}_2$. Again no routing is considered and it is supposed that each packet has a journey that consists

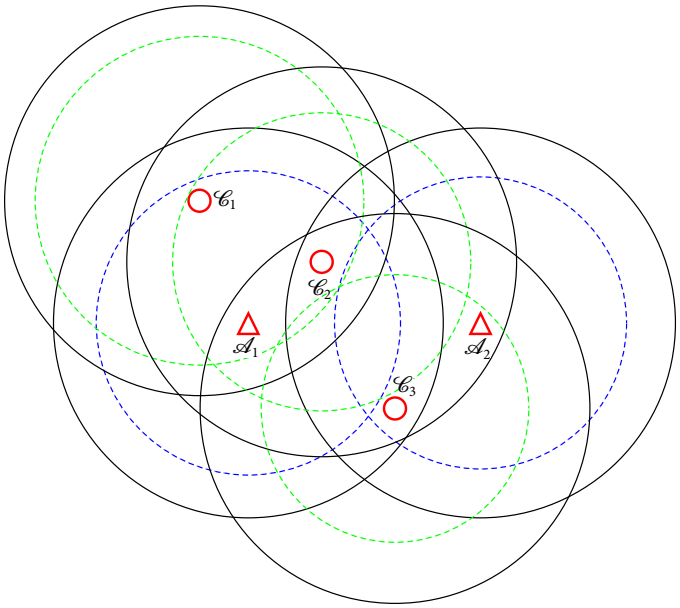


FIGURE 6.10. Simple topology 1 with CCA threshold adaptation. Large black solid circles surrounding each node represent legacy CCA threshold or TP. Blue and green dashed circles represent reduced CCA/TP for APs and clients respectively.

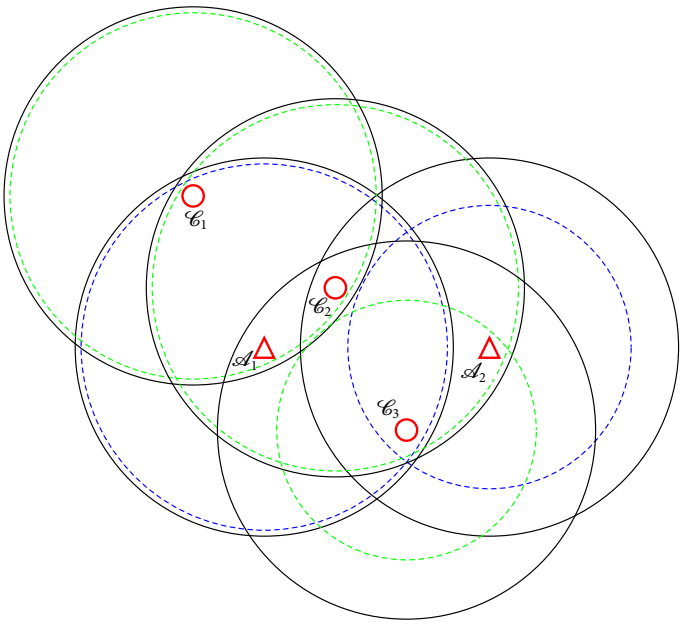


FIGURE 6.11. Simple topology 2 with CCA threshold adaptation. Large black solid circles surrounding each node represent legacy CCA threshold or TP. Blue and green dashed circles represent reduced CCA/TP for APs and clients respectively. $\mathscr{C}_1$ is slightly further away from $\mathscr{A}_1$ than in Fig. 6.10, therefore reducing Cell 1's ability to adapt CCA/TP.

of a single hop. Thus, one may prescribe demand d_1, d_2, d_3 for each of the three links respectively.

Hence, as for the random network example, a collision matrix **E** (see Fig. 6.12) and knowledge of the network matrix **F** (see Fig. 6.13) are used to describe each operational scenario, and these are now just three-by-three in size.

- At legacy CCA threshold and TP values, $\mathscr{C}_1$ can carrier sense transmissions from $\mathscr{A}_1$ and $\mathscr{C}_2$; $\mathscr{C}_2$ can carrier sense transmissions from $\mathscr{C}_1$ and $\mathscr{C}_3$ and $\mathscr{A}_1$ and $\mathscr{A}_2$; $\mathscr{C}_3$ can carrier sense transmissions from $\mathscr{C}_2$, $\mathscr{A}_1$ and $\mathscr{A}_2$; $\mathscr{A}_1$ can carrier sense transmissions from $\mathscr{C}_1$, $\mathscr{C}_2$ and $\mathscr{C}_3$; $\mathscr{A}_2$ can carrier sense transmissions from $\mathscr{C}_2$ and $\mathscr{C}_3$.

- $\mathscr{C}_1$ and $\mathscr{C}_2$ transmitting simultaneously, at legacy conditions, will result in a collision causing both transmissions to fail. A collision will occur as both clients have the same destination AP which is only capable of receiving one transmission at any one time. Due to the received signal strength from both clients at the AP being similar, both mutually cause interference too great to allow the other transmission to be heard.

- $\mathscr{C}_2$ and $\mathscr{C}_3$ transmitting simultaneously, at legacy conditions, will result in a collision causing both transmissions to fail. Despite the destinations of $\mathscr{C}_2$ and $\mathscr{C}_3$ transmissions being different ($\mathscr{A}_1$ and $\mathscr{A}_2$ respectively), the interference received at each AP (say $\mathscr{A}_1$) from the client whose destination is not that AP ($\mathscr{C}_2$) is roughly similar to the received signal from the client whose intended destination is that AP ($\mathscr{C}_1$). The resulting SINR of the transmission intended for that AP, has too high an interference level for the AP to interpret the message.

- Prior to any optimisation, due to the position of $\mathscr{C}_1$ and $\mathscr{C}_3$ on opposite sides of $\mathscr{A}_1$ out of carrier sensing range of each other, a hidden node issue exists [126] at $\mathscr{A}_1$. $\mathscr{C}_1$ and $\mathscr{C}_3$ transmitting simultaneously, at legacy conditions, will result in a collision causing the transmission from $\mathscr{C}_1$ only to fail. Despite the destination of $\mathscr{C}_1$ and $\mathscr{C}_3$ transmissions being different ($\mathscr{A}_1$ and $\mathscr{A}_2$ respectively) the interference received at $\mathscr{A}_1$ from $\mathscr{C}_3$ is roughly similar to the received signal from $\mathscr{C}_1$. This results in a SINR at $\mathscr{A}_1$ where the interference of $\mathscr{C}_3$ is too high for $\mathscr{C}_1$'s transmission to be received. $\mathscr{A}_2$ cannot hear the transmissions of $\mathscr{C}_1$ and so the transmissions of $\mathscr{C}_3$ are not interrupted. This asymmetry is reflected in matrix **E** in Fig. 6.12(a), i.e., $E_{13} = 1$ but $E_{31} = 0$. The hidden node effect is captured by the difference in the matrices **E** and **F** (legacy CCA values) of Figs. 6.12(a) and 6.13(a) respectively showing that links do not have awareness of a potential collision, i.e., $E_{13} = 1$ but $F_{13} = 0$.

- Clients and APs would not knowingly attempt to reduce their CCA sensitivity to become deaf to other nodes in their cell as this would create a hidden node issue [126] which is known to have a negative impact on performance. Similarly, when optimising transmission

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$ (a)

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$ (b)

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix}$$ (c)

FIGURE 6.12. Collision matrices **E**.

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$ (a)

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$ (b)

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$ (c)

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$ (d)

FIGURE 6.13. Knowledge of the network matrices **F**.

power, nodes would not reduce their transmission power such that others in their cell could not hear them, as this too would lead to a hidden node issue.

The same adaptation techniques applied to the randomly generated networks are applied to the simple smaller networks. The operational scenarios are described in table 6.3.1 with matrices **E** and **F** show in Figs. 6.12 and 6.13 respectively.

6.3.1 Simple Topology Operational Scenarios

All possible optimisations using CCA threshold, TP adaptation or a combination of the two capable of impacting the link structure of the simple networks are identified (Figs. 6.10 and 6.11), i.e., adaptations that will change the link interference and / or link carrier sensing ability from those of the legacy CCA threshold and TP. Tab. 6.3.1 summarises the legacy and then eight possible combinations of optimisation using CCA threshold and TP adaptation. The table explains if and how these can be captured in terms of matrices **E** and **F** shown in Figs. 6.12 and 6.13 respectively.

When applying adaptation techniques (CCA, TP or combined), in some instances the resulting **E** and **F** matrix are the same for both of the simple network topologies (i.e., Numbers 1, 4 and 7 in Tab. 6.3.1). For others, it is only possible to apply the optimisation to one topology (i.e., Numbers 3, 6, 8 and 9 in Tab. 6.3.1). In other cases, applying different optimisation methods result in the same **E** and **F** (i.e., Number 4 topology 1 & 2 and Number 2 topology 2, Number 2 topology 1 and Number 5 topology 1, Number 5 topology 2 and Number 7 topology 1 & 2 in Tab. 6.3.1). Only eight possible unique combinations of **E** and **F** (in Figs. 6.12 and 6.13) exist, capturing all possible optimisation combinations described in Sec. 6.3 and summarised in Tab. 6.3.1. These are labelled (S1-S8) in Tab. 6.3.1.

Table 6.3.1 Operational Scenarios

No.	Description	Topology 1 (Fig. 6.10)	Topology 2 (Fig. 6.11)
1	Legacy CCA Legacy Transmission Power	**E** = Fig. 6.12(a) **F** = Fig. 6.13(a) **(S1)**	**E** = Fig. 6.12(a) **F** = Fig. 6.13(a) **(S1)**
2	Max allowable Reduced CCA Sensitivity all nodes Legacy Transmission Power	**E** = Fig. 6.12(a) **F** = Fig. 6.13(b) **(S2)**	**E** = Fig. 6.12(a) **F** = Fig. 6.13(c) **(S4)**
3	Reduced CCA Sensitivity $\mathscr{C}_1$, $\mathscr{C}_2$ and $\mathscr{A}_1$ only Legacy Transmission Power	**E** = Fig. 6.12(a) **F** = Fig. 6.13(d) **(S3)**	$\mathscr{C}_2$ or $\mathscr{A}_1$ cannot reduce CCA and still hear all nodes in its cell
4	Reduced CCA Sensitivity $\mathscr{C}_3$ and $\mathscr{A}_2$ only Legacy Transmission Power	**E** = Fig. 6.12(a) **F** = Fig. 6.13(c) **(S4)**	**E** = Fig. 6.12(a) **F** = Fig. 6.13(c) **(S4)**
5	Legacy CCA Max allowable Reduced Transmission Power all nodes	**E** = Fig. 6.12(a) **F** = Fig. 6.13(b) **(S2)**	**E** = Fig. 6.12(b) **F** = Fig. 6.13(d) **(S6)**
6	Legacy CCA Reduced Transmission Power $\mathscr{C}_1$, $\mathscr{C}_2$ and $\mathscr{A}_1$ only	**E** = Fig. 6.12(c) **F** = Fig. 6.13(c) **(S5)**	$\mathscr{C}_2$ cannot reduce transmission power and still be heard by all nodes in its cell
7	Legacy CCA Reduced Transmission Power $\mathscr{C}_3$ and $\mathscr{A}_2$ only	**E** = Fig. 6.12(b) **F** = Fig. 6.13(d) **(S6)**	**E** = Fig. 6.12(b) **F** = Fig. 6.13(d) **(S6)**
8	Reduced CCA Sensitivity $\mathscr{C}_1$, $\mathscr{C}_2$ and $\mathscr{A}_1$ only Reduced Transmission Power $\mathscr{C}_3$ and $\mathscr{A}_2$ only	**E** = Fig. 6.12(b) **F** = Fig. 6.13(b) **(S7)**	$\mathscr{C}_2$ or $\mathscr{A}_1$ cannot reduce CCA and still here all nodes in its cell
9	Reduced CCA Sensitivity $\mathscr{C}_3$ and $\mathscr{A}_2$ only Reduced Transmission Power $\mathscr{C}_1$, $\mathscr{C}_2$ and $\mathscr{A}_1$ only	**E** = Fig. 6.12(c) **F** = Fig. 6.13(d) **(S8)**	C_2 cannot reduce transmission power and still be heard by all nodes in its cell

6.3.2 Simple Topology Simulation Methodology

The investigation simulates each of the Scenarios S1-S8 (see Tab. 6.3.1) using the conceptual simulation model as Sec. 6.2 and Chapter 4 (again $Countdown_{\max} = 2$). Due to the simple examples' significantly reduced parameter space, for each three-link scenario, rather than focus on fixed equal demands, a large ensemble of simulations are performed, each with a different demand vector (d_1, d_2, d_3). Specifically, as before, let d_{sat} be the maximum transmission rate of a single client transmitting alone in a clear channel. Each of the demands d_1, d_2 and d_3 is varied independently from 0 to d_{sat} in 40 equal increments, resulting in $41^3 = 68,921$ simulations for each of the eight scenarios. The duration of each simulation is set at $10,000/d_{\text{sat}}$ (i.e., the time needed to transmit 10,000 packets without collisions).

For each individual simulation, statistics on the total number of packets successfully transmitted, the time evolution of queues and the average latency per packet are gathered. These statistics are post-processed with various heuristics (see Sec. 4.3) to decide whether each simulation is within the capacity region, meaning the network and protocol can meet the prescribed demand; or outside the capacity region, meaning that in the large time limit, latency and at least one queue grows without bound.

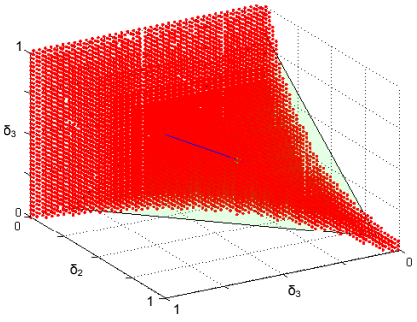
The results that follow (Sec. 6.3.3) are based on comparisons between Scenarios S1-S8. The proportion of simulated demand vectors that are within the capacity region and asymmetry of those vectors provide interesting insight.

6.3.3 Simple Topology Results

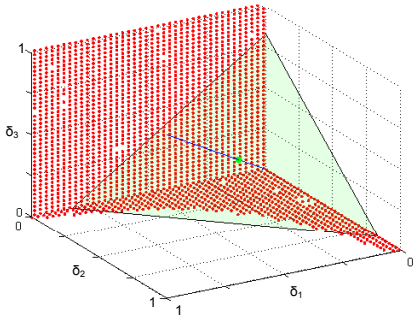
For each of the eight scenarios, the simulation results can be presented in the form of a three-dimensional scatter plot, see Fig. 6.14, where red markers indicate combinations of demand that are within the capacity region. In this plot $\delta_1 := d_1/d_{\text{sat}}$, $\delta_2 := d_2/d_{\text{sat}}$ and $\delta_3 := d_3/d_{\text{sat}}$ denote non-dimensional demand intensities that range from 0 to 1. The concavity of the capacity region (the volume covered by red dots) is apparent from such plots and represents the loss of efficiency in the channel due to competition between transmitters and the resulting collisions. As discussed in previous chapters (see Secs. 4.3, 5.2.4 and 5.3.5), numerical measures that can be derived from these plots are required in order to compare the scenarios.

Firstly, consider triangular cross sections through Fig. 6.14 of total demand intensity $\delta = \text{const}$. The dependence of the proportion of the corresponding triangular area that is within the capacity region S on δ may then be studied and compared across the eight scenarios: see Fig. 6.15(a). Secondly, one may identify the maximum value δ_{cap} (of $\delta_1 = \delta_2 = \delta_3$) which is within the capacity region, and further, one may measure the proportion of the volume (V) $[0,1] \times [0,1] \times [0,1]$ that is within the capacity region — and compare across scenarios. See the caption under each sub figure in Fig. 6.14.

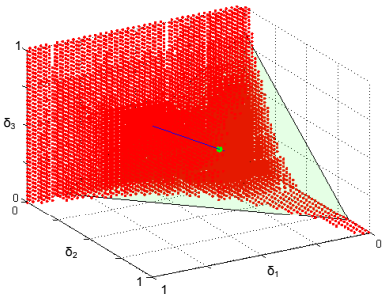
It is additionally possible to consider the demand combination $\delta_1 = \delta_2 = \delta_3 = \delta_{\text{sat}}$ and identify the proportion of that demand met by each of the three respective links δ_{met}. See the caption



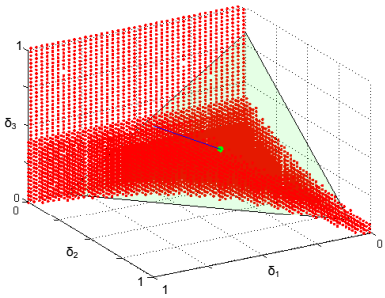
(a) S1, $V = 0.182$, $\delta_{\text{cap}} = 0.34$, $\delta_{\text{met}} = [0.501, 0.004, 0.490]$, ΔF=0, ΔE=0



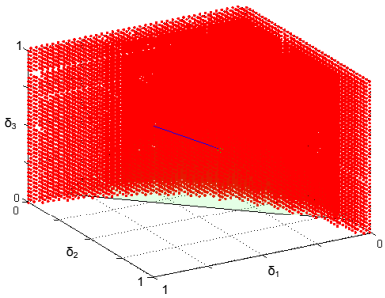
(b) S2, $V = 0.031$, $\delta_{\text{cap}} = 0.02$, $\delta_{\text{met}} = [0.214, 0.044, 0.129]$, ΔF=0, ΔE=2



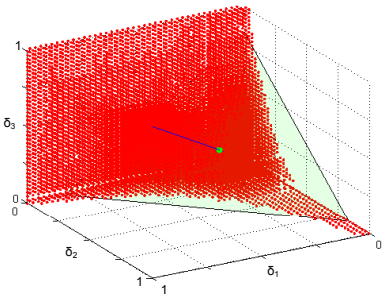
(c) S3, $V = 0.138$, $\delta_{\text{cap}} = 0.27$, $\delta_{\text{met}} = [0.213, 0.046, 0.132]$, ΔF=0, ΔE=1



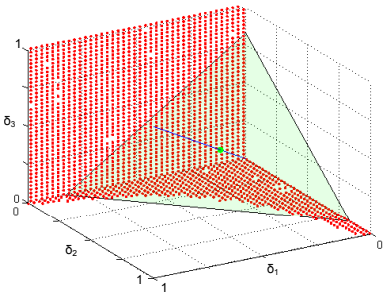
(d) S4, $V = 0.105$, $\delta_{\text{cap}} = 0.27$, $\delta_{\text{met}} = [0.492, 0.007, 0.490]$, ΔF=0, ΔE=1



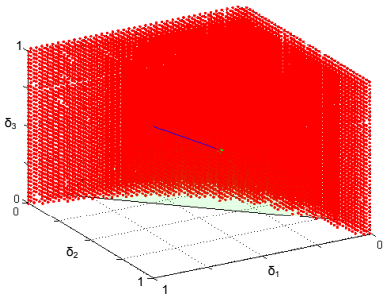
(e) S5, $V = 0.318$, $\delta_{\text{cap}} = 0.32$, $\delta_{\text{met}} = [0.492, 0.014, 0.497]$, ΔF=1, ΔE=1



(f) S6, $V = 0.139$, $\delta_{\text{cap}} = 0.22$, $\delta_{\text{met}} = [0.215, 0.046, 0.132]$, ΔF=2, ΔE=1

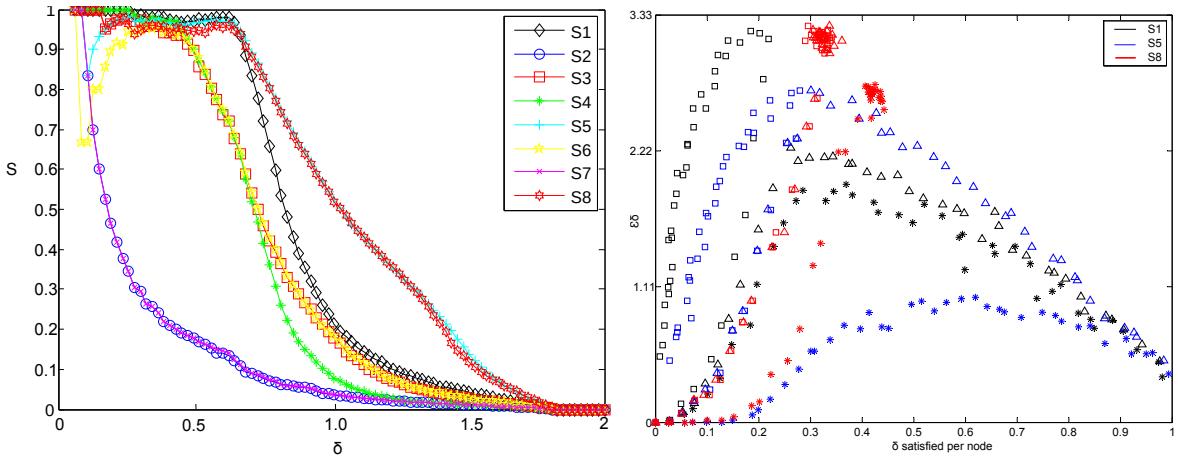


(g) S7, $V = 0.031$, $\delta_{\text{cap}} = 0.02$, $\delta_{\text{met}} = [0.213, 0.044, 0.128]$, ΔF=2, ΔE=2



(h) S8, $V = 0.313$, $\delta_{\text{cap}} = 0.34$, $\delta_{\text{met}} = [0.166, 0.154, 0.212]$, ΔF=1, ΔE=2

FIGURE 6.14. Capacity region plots for S1-S8. Red dots indicate a point is inside the capacity region. Here V is the volume of the capacity region and δ_{cap} is its extent along the line $\delta_1 = \delta_2 = \delta_3$. Further δ_{met} identifies the proportion of the demand combination $\delta_1 = \delta_2 = \delta_3 = \delta_{\text{sat}}$ met by each link respectively. ΔE and ΔF are the reduction in the total number of unit entries relative to legacy S1 in matrices $\mathbf{E}$ and $\mathbf{F}$ respectively, i.e., colliding link combinations and suppressed transmissions respectively.



(a) (b)

FIGURE 6.15. (a) Proportions S of triangular cross section within the capacity region, as a function of total demand intensity $\delta = \delta_1 + \delta_2 + \delta_3$, compared across eight scenarios. (b) A plot of demand (δ) per client against $\varepsilon\delta$ for scenarios S1 (legacy CCA and TP), S5 (legacy CCA and reduced TP to $\mathscr{C}_1$, $\mathscr{C}_2$ and $\mathscr{A}_1$ only) and S8 (reduced CCA to $\mathscr{C}_3$ and $\mathscr{A}_2$ only and reduced TP to $\mathscr{C}_1$, $\mathscr{C}_2$ and $\mathscr{A}_1$ only). Here ε is the failure rate, i.e., the number of failed transmissions per successful transmission, along the line $\delta_1 = \delta_2 = \delta_3$. The triangle, square and star shapes represent links $\mathscr{L}_1$, $\mathscr{L}_2$ and $\mathscr{L}_3$ respectively. The colours indicate the scenario (S1) black, (S5) blue (S8) red.

under each subfigure in Fig. 6.14.

Further, for selected scenarios, the ratio of satisfied demand to failed transmissions is examined (on the line $\delta_1 = \delta_2 = \delta_3$). See Fig. 6.15(b). From this figure the demand (δ) at which each link reaches saturation is apparent. Further insights into the competition for limited channel access between the links can be observed.

The point of the various measures is that they allow one to distinguish whether capacity is added by enhancing goodput in symmetric demand situations, or by allowing fresh combinations of highly asymmetric flows.

6.3.4 Simple Topology Analysis

S5 and S8 were the only scenarios simulated to show any gain in capacity region volume (V) from S1 with legacy CCA threshold and TP. Both of these scenarios are defined by the same **E** matrix in Fig. 6.12(c) and their **F** matrices are shown in Figs. 6.13(c) and 6.13(d) respectively, both containing a hidden node. Due to the adaptation to the TP and CCA threshold in both S5 and S8, the link connecting $\mathscr{A}_2$ with $\mathscr{C}_3$ has a reduced risk of failure due to collision, because the TP in the neighbouring cell has been reduced such that a transmission from $\mathscr{C}_3$ to $\mathscr{A}_2$ (or vice-versa) will always be heard over other interference from the network. Scenario S5 showed

the greatest increase in V improving from $V = 0.182$ for S1 to $V = 0.318$ for S5, narrowly ahead of $V = 0.313$ for S8. Notably for S5, δ_{cap} showed a slight decrease from S1's $\delta_{\text{cap}} = 0.34$ to $\delta_{\text{cap}} = 0.32$ whereas S8 remained the same as S1.

CCA and TP adaptation can be beneficial and allow simultaneous transmissions that previously would have been prevented as illustrated in Fig. 6.1. In the networks investigated in Figs. 6.10 and 6.11, a further problem becomes apparent in that adjusting nodes' ability to hear each other can increase hidden node effects. It is expected that hidden nodes have a significant negative impact to a networks' throughput [126, 127]. The impact of hidden nodes on simple networks was explored in Chapter 5. The asymmetry of the network and the number of hidden nodes (i.e., where the corresponding position in matrix $\mathbf{E}$ shows a '1' but $\mathbf{F}$ shows a '0') however does not directly impact on the volume of the capacity region or δ_{met}, rather the position and asymmetry of the hidden node within the network has a far greater significance. Scenarios S2, S3, S4, S6 and S7 all showed a reduction in V and δ_{cap}. In S5 and S8 the reduced TP to Cell 1 creates a hidden node effect in that Cell 2 can now not hear the transmissions of Cell 1 but Cell 1 cannot hear Cell 2. This however, increases the overall capacity region volume, by eliminating any competition for channel access for the link in Cell 2, facilitating it in meeting its demand vector at certain asymmetric demand combinations. There is no correlation between the number of hidden nodes or asymmetry of the network and total capacity achieved.

It is apparent that δ_{met} does not necessarily increase with V (see Fig. 6.14). Most notably, although showing similar gains in V, scenarios S5 and S8 achieve very different δ_{met}, with the total of S8 approximately half that of S5 and a significant reduction from S1. S5 was the only scenario to show an increased total δ_{met} from S1. Fig. 6.15(b) provides insight into the competition between links considering equal demands. For the three scenarios displayed, the legacy (S1) and two scenarios that showed an increase in V (S5 and S8); the plot shows how the links behave for equal demands. S8 (red markers Fig. 6.15(b)) shows the error rate increases with increased demand to the three links. The clustering of three plotted red shapes representing the three links shows how all links saturate at similar ratios of demand to error. S1 and S5 show a different trend. Initially, as S8, the error rate of S1 and S5 links increases with demand, however, both reach a point where as further demand is added, the error rates for links $\mathscr{L}_1$ and $\mathscr{L}_3$ decrease. Beyond the same point, the ability of link $\mathscr{L}_2$ to meet its demand decreases and further its error rate also decreases. This indicates that the link has a lower probability of accessing the channel, losing out in competition for the limited resource to the other two links. This starvation effect [93, 127] can be explained by link $\mathscr{L}_2$'s position between the other two links in the topology. The difference between S5 and S8 can be identified as due to the difference between their knowledge of the network, i.e., Fig. 6.13(c) and 6.13(d) respectively. The matrices contain the same number of 1s, however, the positioning indicates nodes' different knowledge of the network surrounding them and this results in a significantly higher total δ_{met} for S5.

6.4 Discussion

The results of this study relating to CCA adaptation support existing work and are comparable in terms of the magnitude of performance gain. Paper [5] found CCA adaptation could achieve up to $\approx 20\%$ throughput improvement if implemented appropriately and further could improve fairness. As with this study, [5] recognises the potential improvement to be highly dependent on topology. The TP adaptation results are comparable to a study [176] carried out using the OPNET [199] simulator. That study was able to show throughput improvements of several hundred percent in certain situations.

Through further examining modified CCA thresholds in simple networks, this study was able to identify specific changes to performance stemming from the underlying 'knowledge of the network' matrix **F** in the model. Considering the topology in Fig. 6.10, Cell 1 nodes ($\mathscr{C}_1$ and $\mathscr{C}_2$ along with their associated $\mathscr{A}_1$) can all reduce their CCA sensitivity such that they can only hear each other and are deaf to the transmissions of Cell 2 nodes ($\mathscr{C}_3$ and $\mathscr{A}_2$). Similarly, Cell 2 nodes can do the same making themselves deaf to the transmissions of $\mathscr{C}_1$, $\mathscr{C}_2$ and $\mathscr{A}_1$. Consider the topology in Fig. 6.11; because of the slight change of position of $\mathscr{C}_1$, in order for Cell 1 nodes to all hear each others' transmissions, they cannot adapt their CCA threshold by a significant margin. This simultaneously has the effect of preventing $\mathscr{A}_1$ and $\mathscr{C}_2$ reducing CCA sensitivity to become deaf to $\mathscr{C}_3$. Cell 2's nodes are unaffected by the position of $\mathscr{C}_1$ and so can reduce their CCA sensitivity to become deaf to the other three nodes in the network. This results in the matrix **F** becoming asymmetric when CCA sensitivity is reduced, such that $\mathscr{C}_2$ and $\mathscr{A}_1$ can hear nodes in their neighbouring cell (i.e., $\mathscr{A}_2$ and $\mathscr{C}_3$) but $\mathscr{C}_3$ and AP_2 are deaf to their neighbours.

Adapting the TP for any number of clients showed the largest range of performance change, potentially leading to the greatest throughput gain or the most significant loss. It was the adaptation that most consistently led to an increase in throughput performance. Whereas modifying the CCA threshold can only enable pairs of links to simultaneous transmit that would previously have been silenced (i.e., change matrix **F**), modifying the TP additionally has the potential to enable pairs of links that would have previously collided when transmitting simultaneously (i.e., changing matrix **E** also). When all clients reduced their TP by the same proportion, the SINR at the access points remains the same and therefore matrix **E** stayed the same. In this case, reducing the volume of transmission from all nodes has much the same effect as reducing the CCA sensitivity at all nodes changing matrix **F**. More significant changes occurred when the transmission power is adjusted asymmetrically.

Adaptation of TP is subject to similar topological constraints as discussed with CCA adaptation. In the first simple network (Fig. 6.10) it is possible for Cell 1 to reduce TP where Cell 2 cannot hear (but still the nodes in Cell 1 can hear each other). However, in the second network (Fig. 6.11), $\mathscr{C}_2$ cannot reduce TP to a level where Cell 2 nodes cannot hear whilst still being heard by all nodes in its cell. When $\mathscr{C}_2$ reduces its TP, the SINR ratio at $\mathscr{A}_2$ changes such that when $\mathscr{C}_2$ and $\mathscr{C}_3$ transmit simultaneously the louder transmission from $\mathscr{C}_3$ is always heard over

$\mathscr{C}_2$ (i.e., $E_{32} = 0$). There is still of course interference at $\mathscr{A}_1$, such that in the case of $\mathscr{C}_2$ and $\mathscr{C}_3$ transmitting simultaneously, $\mathscr{A}_1$ cannot interpret its signal from $\mathscr{C}_2$ and the transmission always fails. The case where Cell 2 only reduces TP to the minimum value required for $\mathscr{A}_2$ to hear $\mathscr{C}_3$ and vice-versa removed the risk of a potential collision between $\mathscr{C}_3$ and $\mathscr{C}_1$ at $\mathscr{A}_1$ (i.e., $E_{13} = 0$). Further, now the SINR ratio at $\mathscr{A}_1$, when $\mathscr{C}_2$ and $\mathscr{C}_3$ transmit simultaneously, is modified such that the louder transmission from $\mathscr{C}_2$ is heard over $\mathscr{C}_3$ (i.e., $E_{23} = 0$). Of course interference at $\mathscr{A}_2$ occurred such that when $\mathscr{C}_2$ and $\mathscr{C}_3$ transmit simultaneously, $\mathscr{A}_2$ was unable to receive its signal from $\mathscr{C}_3$.

Reducing the TP at one node effectively generated the same impact as reducing the CCA sensitivity of surrounding nodes in terms of allowing simultaneous transmissions that previously would have been silenced (i.e., by changing **F**). When $\mathscr{C}_3$ reduced its TP such that only $\mathscr{A}_2$ can just hear it, $\mathscr{A}_1$ and $\mathscr{C}_2$ could not hear it with their fixed CCA legacy threshold. Hence matrix **F** changed to reflect this (Fig. 6.13(d)). Likewise when $\mathscr{C}_2$ reduced its TP (first topology only Fig. 6.10) such that only $\mathscr{A}_1$ and $\mathscr{C}_1$ could just hear it, $\mathscr{A}_2$ and $\mathscr{C}_3$ could no longer hear it with their fixed CCA legacy threshold. Hence, matrix **F** changed to reflect this (Fig. 6.13(c)). Because of its capability to impact both matrices **E** and **F**, adapting TP has a far more significant impact than adapting CCA threshold.

The investigation tested randomly applying CCA or TP adaptation to cells in the randomly generated network or combining both within each cell. Random deployment is representative of networks where there is no co-ordination between separately deployed cells. As the results of the randomly generated networks demonstrated (Fig. 6.3), this generally provided a performance improvement, however only slight. In the simple examples, both methods impacted on the cells' ability to hear each other, changing matrix **F**, and collisions, changing matrix **E**. The results in Fig. 6.14(g) and 6.14(h), however, show very different impacts on the capacity of the network depending on which cell reduced CCA sensitivity or TP. This highlights the sensitivity of the adaptation's impact to topology.

In the simple networks (Fig. 6.10 and Fig. 6.11), the options for combining the CCA and TP adaptation are very limited. Simultaneously reducing CCA sensitivity of the receiver and TP of the sender risks breaking the connection between the two nodes. Of course in some situations, a balance could be reached where both transmitter and receiver are able to reduce the TP and CCA sensitivity by a margin and still maintain a connection. One might expect this may minimise their aggregate interference on the surrounding network. The simulation of this on the randomly generated networks (i.e., R4) showed such an adaptation to increase mean throughput performance, however, the change led to a loss on a higher number of occasions than a gain. This indicates the mean is raised by a few outlier networks even though for the majority of networks, the adaptation results in a slight throughput loss.

Each node, when adjusting parameters in accordance with the prescribed rules, sets its TP and / or CCA value such that they can still carrier sense, or be carrier sensed by, the same

nodes of their cell as at legacy values. In practice, this adjustment might be achieved by each node monitoring the received signal strength indication (RSSI) of all of its associated nodes, and by choosing the RSSI of the farthest and setting CCA and TP accordingly [149]. However, in a dynamically changing network, the challenge is more difficult. If a new client joins and associates with an AP, other clients in the same cell, with reduced CCA, might not carrier sense the new client and the new client may not carrier sense those existing clients with reduced TP. All clients associated with a particular cell receive a beacon frame from their respective AP at every beacon interval (typically $\approx$ 100ms) [5]. Beacons distribute information of configuration changes by notifying existing clients of new arrivals to the cell. Hence, if a client who has reduced CCA cannot carrier sense the new client, it may return to legacy settings and re-apply the rule. Alternatively, if a client of reduced TP is made aware of a new client, it may alter its TP to incorporate the new arrival in accordance with the rule.

Other authors have explored individual nodes dynamically optimising CCA and TP parameters [5, 56, 149, 248], however, in any non-trivial network, it is unlikely a true system optimal throughput will be reached by such an approach. Even if nodes cooperated, searching the vast parameter space for an optimal solution will likely not be computationally tractable. Further, these approaches often result in poor fairness. The advantage of the simple rules proposed is that they restrict the parameter space, reducing the required computation, and hence might be applicable to dynamic networks.

Additionally, a relationship between the clustering coefficient and the likely throughput has been identified: the potential performance improvement that adaptations can provide decreases with increasing clustering coefficient. This relationship can be explained as follows. A higher clustering coefficient indicates that the clients are grouped together tightly around the APs with fewer clients falling on the border between cells. When there are few or no clients on the border between cells, cross cell interference cannot be improved by adapting the CCA threshold or TP. The clustering coefficient was used to determine when an adaptation should be applied and showed this to reduce the number of implementations leading to a loss in throughput, and thus to an increase in the mean throughput change. If a network's topology can be monitored and the clustering coefficient computed in real time, rules could be implemented accordingly to achieve maximum gain. Exploring different combinations of parameters from the five rules proposed, and better use of the clustering coefficient and other network characteristics as a measure to predict if and when to apply them, are topics for future research.

As many other studies in related areas do, this study used Jain's Fairness Index [125] which rates the fairness of a set of n links' (with unique transmitters) throughput values where x_i is the throughput for the i^{th} connection. The metric ranges from a minimum of $1/n$ to a maximum of 1 (when all users receive the same allocation). This index is k/n when k users equally share the resource, and the other $n-k$ users receive zero allocation. The fairness impact from the simulations on randomly generated networks was reported in Fig. 6.5. In the simple examples,

the shape of the capacity regions (Figs. 6.14) indicates how some highly asymmetric demand combinations, and therefore asymmetric (i.e., unfair) throughput combinations, can be met by the system where even demand combinations of equivalent total throughput, cannot. The change in results from scenarios S1 to S5 and S8, which both showed an increase in overall volume of the capacity region, show a significant change in fairness. In these two scenarios, the link connecting $\mathscr{C}_3$ and $\mathscr{A}_2$ has a higher chance of transmitting, after the adaptation, than the other two links. This change in fairness, however, is slightly misleading as only the ability to meet the capacity of one link is changed. The probability of the links connecting the two clients to $\mathscr{A}_1$ transmitting successfully remains the same for all three scenarios: the third link's probability of successful transmission is increased without impacting on these. The two scenarios S5 and S8 that show growth in the capacity region to the legacy S1 do so by enabling additional asymmetric demand combinations to be met by the system. Comparing S1, S5 and S8 at δ_{sat}, S8 achieves the fairest throughput with similar values of δ_{met} to all clients, however it also achieves the lowest total δ_{met}. Fairness, although still a useful measure, must be considered carefully. If a network of three transmitting links, each achieving roughly equal channel access (i.e., fair) is adapted, via TP or CCA adaptation, such that two of those three links achieve the same throughput as before but one now achieves greater, the system is now less fair although the only impact to throughput has been positive.

6.5 Conclusion

This chapter contributes to knowledge in the following ways:

CTK 6.1 Presenting a modelling methodology for investigating CCA/TP adaptation.

CTK 6.2 Demonstrating heuristic CCA and TP adaptation methods that result in performance improvements.

CTK 6.3 Identifying that network throughput does not correlate with the number of hidden or exposed nodes when the network contains more than one cell and explaining how asymmetrically suppressing part of a network can remove competition for channel access and lead to overall network throughput improvement.

CTK 6.4 Identifying a relationship between throughput improvement and network clustering coefficient and thus network conditions under which CCA and TP adaptations are most likely to realise throughput improvements.

This chapter explored CCA threshold and TP adaptation considering practical methods by which they could be implemented. CCA, TP and combined CCA/TP adaptation are all capable of generating a performance improvement in WLANs, however, the impact of the adaptation is highly dependent on topology. TP adaptation provides the best chance of a performance

improvement. When applied only to networks with low clustering coefficient, the probability of achieving a performance gain from TP adaptation is increased.

CHAPTER 7

Effects of Routing on the Capacity of Multi-Hop Wireless Networks

This chapter is based on: W. Jones and R. E. Wilson "Effects of Routing on the Capacity of Multi-Hop Wireless Networks," In Proceedings of the 21st ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWIM '18). ACM, New York, NY, USA, 155-162.

This chapter considers multi-hop wireless mesh networks and examines whether capacity may be improved by distributing the data flows for each origin-destination (OD) pair across multiple routes. As in Chapters 5 and 6, the network geometry and the application of technologies, such as beamforming or full-duplex, are described by a matrix that describes which pairs of transmissions (or 'links') are compatible in that they can occur simultaneously without collisions or other conflicts. Under a binary interference framework, the conflict matrix is used to derive the maximal sets of compatible links, and secondly, these are used to derive a system of linear inequalities that bound links' data flows. These steps are computationally expensive, but it is shown how the system usually collapses when re-expressed in terms of flows on routes. The theory is illustrated in terms of a simple 'Braess' network example, with a single OD pair, and compared with simulation results. Then networks with two OD pairs, whose data flows have some nodes in common and thus contend with each other, are considered. It is shown how networks can be designed such that they exploit multiple relay nodes and routes so as to increase capacity. Further, the same problem is examined on ensembles of random networks. It was found that in many cases, capacity can be improved if the OD pairs distribute their traffic over several routes. The chapter poses and solves a set of linear programs that model (i) cooperative behaviour; (ii) the optimisation of one OD pair when presented with a fixed route

assignment by the other; and (iii) variants of these games when both OD pairs are in contention with background (single-hop) traffic.

7.1 Background

Wireless multi-hop networks are of interest to communications providers because of their ability to extend the coverage of a network beyond existing infrastructure, avoiding the need for costly deployment of cables [105, 227, 231]. One proposal is to use multi-hop wireless architectures for backhaul [187]. Generally speaking, wireless multi-hop networks are easy to deploy and can be expanded incrementally. The advantage, in theory, is that they can adjust to changing demands or node failures, by re-routing data flows. They have the potential to provide cost-effective service in areas with no available wired connection. Other potential applications [165, 180] include sensor networks, military operations, emergency services, disaster recovery operations and vehicular ad-hoc networks [159].

There are many advantages of multi-hop systems, such as low up-front cost, easy network maintenance, robustness, reliable service coverage, etc [6, 187]. Hence they are currently undergoing rapid commercialisation in scenarios such as home broadband, building automation, free high-speed wireless access in public areas, and connectivity for transport systems [6].

In some situations, through multi-hop communications, the same coverage can be achieved with much lower TP [6]. Communicating via multi-hop allows the distances over which the nodes need to transmit to be reduced. This has the potential to increase network throughput due to lower pathloss and better spatial reuse [187].

Emerging physical-layer technologies, such as directional antennas [208, 260], i.e., 'beamforming', can help reduce interference and increase the capacity of wireless multi-hop networks (see Chapter 5). Alternatively, adapting the CCA threshold or TP [64, 142, 248] or incorporating full-duplex technology [133] can achieve the same outcome (see Chapters 5 and 6). Each of these technologies works by allowing extra combinations of transmissions to occur simultaneously without collisions or other impediment.

Several challenges remain in enabling multi-hop wireless networks to compete with the throughput and delay guarantees of wired backhauls. Multi-hopping has the potential to boost throughput, but the network must now be more carefully designed to mitigate interference and minimise delays. It seems likely that a wireless multi-hop network's robustness will be improved if each origin-destination (OD) pair (i.e., end-to-end pair) exploits several routes across the network. In contrast, the goal of this chapter is to explore how using multiple routes may reduce contention and increase capacity. The difficulty is identifying which routes, and resources on those routes, have spare capacity and hence directing data traffic along them, avoiding others which are overloaded [139]. Modelling these multi-hop networks is a significant challenge in any non-trivial network topology.

The approach of this chapter (following Tassiulas and Ephremides [243, 244]), is to use a continuous flow-level model and to neglect the problem of how a scheduler would achieve the required time division. Moreover, as throughout this thesis, it is supposed all nodes operate on a single channel (no frequency division). Nevertheless, identifying a wireless multi-hop network's capacity region (i.e., the set of achievable data flows) is a significant computational challenge [11] for dense networks (NP-hard in the general case). Thus the work of Gummadi et al. [103] has provided polynomial-time methods for approximating the capacity region of single-hop [103] and multi-hop [104] networks assuming bounds on the interference structure. In contrast, the method presented in this chapter uses an *exact* approach. The surprise is that although the intermediate computational steps are expensive, the resulting constraints are often rather simple when expressed in terms of route flows.

For the methodology that follows in Sec. 7.2, a highly idealised situation is assumed in which loss of efficiency due to multiple links in direct contention is neglected (i.e., not a concave capacity region — see Fig. 3.10 p.76). Similarly, the possibility of additional demands generating a slight gain in capacity for multiple links in direct contention is neglected (i.e., not a convex capacity region — as observed in [151] illustrated in Fig. 3.11 p.77). Further, this analysis is based on constraints due to collisions only and does not consider cases where nodes able to transmit do not due to other transmissions occurring in the network (exposed nodes — see p.58).

Consider a network of n possible transmissions or 'links'. As in previous chapters, a collision matrix $\mathbf{E}$ can be used to describe interactions between links, with entries E_{ij}, $1 \le i,j \le n$, where if $E_{ij} = 1$, then links i and j cannot transmit simultaneously, whereas if $E_{ij} = 0$ then links i and j can transmit simultaneously. Interference (and hence the collision matrix) is not always symmetric: that is, if link i collides with link j, in some cases only one link transmission may fail, i.e., $E_{ij} = 1$ but $E_{ji} = 0$.

Further, the 'knowledge of the network' matrix $\mathbf{F}$ describe each link's awareness of the network around it, with entries $F_{ij}, 1 \le i,j \le n$. If $F_{ij} = 1$, then link i is aware of j and cannot transmit simultaneously, whereas if $F_{ij} = 0$ then link i either can transmit simultaneously or is unaware of j (cannot hear its transmissions). As with the collision matrices, the network topology (and hence the matrix $\mathbf{F}$) is not always symmetric: that is, if link i has knowledge of link j, then link j does not necessarily have knowledge of link i, i.e., $F_{ij} = 1$ but $F_{ji} = 0$. A lack of knowledge of each other's transmissions can result in the clients transmitting simultaneously, each not realising the other is transmitting, causing a collision.

Let x_i, $i = 1,2,...,n$, be the data flow for each link. Flows will be measured as a proportion of the time each transmitter is successfully transmitting, so $0 \le x_i \le 1$ for each i. The goal is to derive the upper bounds on the capacity region, i.e., upper bounds on the set of x_i, that can be successfully supported and thus identify the maximum achievable traffic flow.

If $E_{ij} = 1$, then $x_i + x_j \le 1$, because transmitters i and j cannot successfully transmit simultaneously and so must time-share. The same conclusion holds if $E_{ji} = 1$. So, therefore, although

asymmetry in $\mathbf{E}$ may play a role in the capacity region's concavity, it plays no role in the bounds defined as follows. Hence this analysis assumes $\mathbf{E}$ is symmetric or has been symmetrised. This symmetrised conflict matrix is denoted $\tilde{\mathbf{E}}$.

The chapter proceeds as follows. Sec. 7.2 develops a theoretical framework for computing the capacity region. Sec. 7.3 illustrates the theory for a simple 'Braess' network. Sec. 7.4 provides a comparison of the analytical method with simulation results. Sec. 7.5 examines how networks can be designed to maximise capacity for a situation in which two OD pairs contend with each other. Sec. 7.6 examines the same problem on ensembles of random networks. Finally, Sec. 7.7 discusses the findings and Sec. 7.8 draws conclusions.

7.2 A Theoretical Framework for Computing the Capacity Region

The starting point is a symmetric $n \times n$ matrix $\tilde{\mathbf{E}}$, with diagonal entries $\tilde{E}_{ii} = 0$, that describes the conflict between n possible transmissions or 'links'. If $\tilde{E}_{ij} = \tilde{E}_{ji} = 0$ for $i \neq j$, then links i and j are *compatible*, i.e., may transmit simultaneously without interference or other impediment. If $\tilde{E}_{ij} = \tilde{E}_{ji} = 1$, then links i and j are not compatible, either because i interferes with j, j interferes with i, or because they are transmissions on the same channel from the same node.

Larger sets of compatible links are described by n-vectors $\mathbf{z}$, where

$$z_i = \begin{cases} 1, & \text{if link } i \text{ is in the compatible set,} \\ 0, & \text{otherwise.} \end{cases} \tag{7.1}$$

Because the method employs a binary interference model, a set of links is compatible if and only if every pair in the set is compatible. It follows that

$$\mathbf{z}^{\mathrm{T}} \tilde{\mathbf{E}} \mathbf{z} = \sum_i \sum_j \tilde{E}_{ij} z_i z_j = 0, \tag{7.2}$$

since any positive term in the sum has $\tilde{E}_{ij}$, z_i and z_j equal to one, and thus corresponds to an incompatible pair of links. (Throughout, the method employs column vectors by default and T denotes matrix transpose.)

Since the aim is to find upper bounds on capacity, the method seeks to determine sets of links that are compatible and *maximal* — i.e., if any link were to be added to such a set, it would become incompatible. The corresponding vectors $\mathbf{z}$ are maximal in the sense that no other vector satisfying (7.1,7.2) is superior in the sense of the standard partial ordering on $\mathbb{R}^n$.

Such maximal vectors are found most efficiently by organising the 2^n possibilities into a DAG (directed acyclic graph) with $n+1$ levels $0, 1, 2, \ldots, n$ according to the number of entries equal to one, see Fig. 7.1. A vector in level-k thus has k immediate children in level-$(k-1)$, corresponding to setting in turn each of the parent's k non-zero entries to zero. The search begins

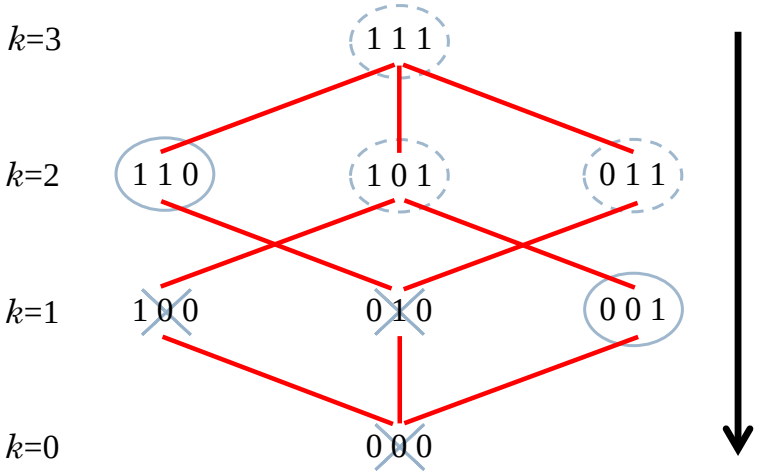


FIGURE 7.1. An illustration of the maximal vector search method for n=3 bits. In this example, level-k has k immediate children in level-$(k-1)$, corresponding to setting in turn each of the parent's non-zero entries to zero. The search begins with the top level (the vector with all entries equal to 1), and proceeds in turn to successively lower levels. When a vector is found that is *maximal*, all its descendants are discarded from the search. In this example, for the level $k=2$ the vector ([1 1 0]) is found to be maximal so its children are discarded from the search; similarly ([0 0 1]) in level $k=1$. Of course, for this simple example, the reduction in the search space is small but for more complicated examples it maybe significant.

with the unique level-n vector $\mathbf{1}_n$ (the n-vector with all entries equal to 1), and proceeds in turn to successively lower levels. When a vector is found that satisfies (7.2), all its descendants are discarded from the search. Supposing p such maximal vectors $\mathbf{z}_i$ are found, these are grouped in the $p \times n$ matrix $\mathbf{Z}$, where $\mathbf{Z}^{\mathrm{T}} = \left(\mathbf{z}_1 | \mathbf{z}_2 | \cdots | \mathbf{z}_p\right)$, so that $Z_{ij} = (\mathbf{z}_i)_j$.

At a system level, capacity is determined by time-sharing amongst the maximal co-transmitting sets. Let α_i denote the proportion of time awarded to $\mathbf{z}_i$. Then

$$\alpha_1 + \alpha_2 + \ldots + \alpha_p \leq 1. \tag{7.3}$$

The problem is to re-express these inequalities in terms of inequalities on the n-vectors $\mathbf{x}$ of flows (data rates) on links. Each link i appears potentially in several different maximal compatible sets — specifically, those j for which $Z_{ji} = 1$. Therefore, for each $i = 1, 2, \ldots, n$,

$$x_i \leq \sum_{j=1}^{p} Z_{ji} \alpha_j. \tag{7.4}$$

Thus the method presented explores different ways of replacing the terms in (7.3) with link flows x_i. Specifically, the approach is to seek n-vectors $\mathbf{c} \geq \mathbf{0}$ that weight equations (7.4), so that

$$\sum_{i=1}^{n} c_i x_i \leq \sum_{i=1}^{n} c_i \left(\sum_{j=1}^{p} z_{ji} \alpha_j \right) = \sum_{j=1}^{p} \left(\sum_{i=1}^{n} c_i z_{ji} \right) \alpha_j. \tag{7.5}$$

Hence provided $\sum_{i=1}^{n} c_i Z_{ji} \leq 1$ for $j = 1,2,\ldots,p$, or equivalently $\mathbf{Z}^{\mathrm{T}}\mathbf{c} \leq \mathbf{1}_p$, the method obtains $\sum_{i=1}^{n} c_i x_i \leq 1$, or equivalently $\mathbf{c}^{\mathrm{T}}\mathbf{x} \leq 1$.

It may be shown that a 'spanning' set of such vectors $\mathbf{c}$ consists of those whose entries are either 0 or 1 and which are maximal in terms of the standard partial ordering on $\mathbb{R}^n$, and thus the procedure to find them is the same DAG-search which established maximal $\mathbf{z}$.

Supposing q such maximal vectors $\mathbf{c}_i$ are found, they can be grouped in the $q \times n$ matrix $\mathbf{C}$, where $\mathbf{C}^{\mathrm{T}} = \left(\mathbf{c}_1 | \mathbf{c}_2 | \ldots | \mathbf{c}_q\right)$, so that $C_{ij} = (\mathbf{c}_i)_j$. It follows that the constraints on link flows may be expressed in the form

$$\mathbf{C}\mathbf{x} \leq \mathbf{1}_q, \quad \text{with} \quad \mathbf{x} \geq \mathbf{0}. \tag{7.6}$$

Suppose now there are r routes under consideration and let $\mathbf{A}$ denote the $r \times n$ route-link incidence matrix, so that $A_{ij} = 1$ if route i uses link j, and 0 otherwise. Let $\mathbf{y}$ denote the r-vector of flows on each route. Hence $\mathbf{x} = \mathbf{A}^{\mathrm{T}}\mathbf{y}$. Thus recast in route flows, (7.6) becomes $(\mathbf{C}\mathbf{A}^{\mathrm{T}})\mathbf{y} \leq \mathbf{1}_q$, with $\mathbf{y} \geq \mathbf{0}$. However, in contrast to (7.6), this system typically has redundancy in that a small subset of the constraints describes the same feasible polytope as the full set. Hence although the intermediate objects $\mathbf{Z}$ and $\mathbf{C}$ can be large and their computation is formally NP-hard, they may be computed (or approximated [104]) off-line and a simpler reduced system be used for on-line route-balancing.

The route-flow system is therefore reduced by testing each constraint with a linear program that maximises it as an objective function with respect to the system of constraints in which it is removed. The resulting $\tilde{q} \leq q$ expresses constraints in the form

$$\tilde{\mathbf{C}}\mathbf{y} \leq \mathbf{1}_{\tilde{q}}, \quad \text{with} \quad \mathbf{y} \geq \mathbf{0}. \tag{7.7}$$

For small networks, the resulting system is simple enough for the constraints to be directly interpretable. Moreover, as is shown shortly, there is a simple relationship between $\tilde{q}$ and optimal routing patterns.

This investigation uses system (7.7) to constrain a variety of linear programs depending on the particular problem in hand. For example, in the case of a single OD pair (Sec. 7.3), the study explores maximising the total $\|\mathbf{y}\|_1$ of route flows with respect to (7.7). In the case of two OD pairs A and B (Sec. 7.5 and 7.6), the route flow vector is partitioned by writing $\mathbf{y}^{\mathrm{T}} = (\mathbf{y}_{\mathrm{A}}^{\mathrm{T}} | \mathbf{y}_{\mathrm{B}}^{\mathrm{T}})$. The study can then

- GAME 1: maximise $\|\mathbf{y}_{\mathrm{A}}\|_1$ (resp. $\|\mathbf{y}_{\mathrm{B}}\|_1$) subject to $\|\mathbf{y}_{\mathrm{B}}\|_1 = d_{\mathrm{B}}$ (resp. $\|\mathbf{y}_{\mathrm{A}}\|_1 = d_{\mathrm{A}}$), with both $\mathbf{y}_{\mathrm{A}}$ and $\mathbf{y}_{\mathrm{B}}$ free — that is maximise the throughput for one OD pair supposing that the throughput of the other is fixed, in a setting where routing is coordinated cooperatively to achieve the best outcome. By playing the symmetric variants of the game across the range $0 \leq d_{\mathrm{A}}, d_{\mathrm{B}} \leq 1$, a Pareto front may be determined for $(\|\mathbf{y}_{\mathrm{A}}\|_1, \|\mathbf{y}_{\mathrm{B}}\|_1)$ that describes the boundary of the capacity region.

- GAME 2: maximise $\|\mathbf{y}\|_1$ subject to $\|\mathbf{y}_\mathrm{A}\|_1 = \|\mathbf{y}_\mathrm{B}\|_1$ with both $\mathbf{y}_\mathrm{A}$ and $\mathbf{y}_\mathrm{B}$ free — that is, maximise the combined throughput with fairly shared resource, under cooperative routing. This game yields a point on GAME 1's Pareto front, which is used (Sec. 7.6) as a single figure of merit.

- GAME 3: maximise $\|\mathbf{y}_\mathrm{A}\|_1$ (resp. $\|\mathbf{y}_\mathrm{B}\|_1$) subject to $\|\mathbf{y}_\mathrm{B}\|_1 = d_\mathrm{B}$ (resp. $\|\mathbf{y}_\mathrm{A}\|_1 = d_\mathrm{A}$) with just $\mathbf{y}_\mathrm{A}$ (resp. $\mathbf{y}_\mathrm{B}$) free whilst the alternate OD pair presents a fixed route assignment — for example, distributing demand equally over its routes. This game illustrates what may be achieved without cooperation between the OD pairs.

In addition, these games may be extended to consider more OD pairs, or indeed analysed in the presence of fixed 'background' single-hop link flows $\mathbf{x}_0$, where the theory is re-developed from (7.6) onwards by writing $\mathbf{x} = \mathbf{x}_0 + \mathbf{A}\mathbf{y}$. In particular, the method used is to employ $\mathbf{x}_0 = \beta\mathbf{1}_n$, where $\beta > 0$ is a constant background load applied equally to every link (see Sec. 7.6). Note such background flow changes the system of reduced constraints.

Each of these linear programs has at least one vertex of their feasible region, defined by (7.7) and the game's equality constraint(s), in their optimal solution set. For example, since GAMEs 1 and 2 incorporate a single equality constraint, such vertices are found by requiring that $r-1$ of the $\tilde{q}+r$ inequality constraints are tight and therefore that at most $\tilde{q}+1$ inequality constraints are slack. It follows that at such optimal vertices, at most $\tilde{q}+1$ route flows are non-zero. Thus $\tilde{q}$ provides a measure for each network of how many routes need to be employed to achieve optimal throughput.

7.3 Single OD Pair Tutorial

To illustrate the theory developed in Sec. 7.2, consider the network shown in Figs. 7.2-7.9. This 'diamond' network is well-known in the road traffic modelling literature, where it displays the Braess paradox [38]. However, here it is used merely as the simplest non-trivial network with a single OD pair. The idea is that data must flow from O to D via either a northern or southern relay node, yielding three possible routes; the northern route 1: links $\mathscr{L}_1, \mathscr{L}_4$; the cross-town route 2: links $\mathscr{L}_1, \mathscr{L}_3, \mathscr{L}_5$; and the southern route 3: links $\mathscr{L}_2, \mathscr{L}_5$). Note links $\mathscr{L}_2$ and $\mathscr{L}_3$ (similarly $\mathscr{L}_4$ and $\mathscr{L}_5$) terminate at the same node, so cannot co-transmit. In contrast, links $\mathscr{L}_1$ and $\mathscr{L}_2$ (similarly $\mathscr{L}_3$ and $\mathscr{L}_4$) originate at the same node, so cannot co-transmit.

Figs. 7.2-7.9 part (a) illustrate the PHY layer interactions, with coloured circles (corresponding to coloured links) illustrating transmission radius. Throughout a key (boxes attached to each link with arrows pointing toward or away from the link) is used to indicate compatible / incompatible link pairs. For the purpose of explanation, consider the box attached to $\mathscr{L}_1$ (red link) with the arrow pointing towards the link in Fig. 7.2 (labelled (i)). The five cells of the box (left to right) relate to the five links of the network (link number corresponding to the position). If a position is filled with a colour, this link's transmission will fail if attempting to co-transmit with the link of

the corresponding colour. In this example $\mathscr{L}_1$'s transmission will fail if it attempts to co-transmit with $\mathscr{L}_3$ (blue), $\mathscr{L}_4$ (yellow) or $\mathscr{L}_5$ (purple) due to inference at the receiver. Additionally, in this example $\mathscr{L}_1$ may not co-transmit with $\mathscr{L}_2$, as both originate from the same node. This is indicated by green hashing of the second cell. The first cell is left blank as $\mathscr{L}_1$ may not collide with itself.

Now consider the box attached to $\mathscr{L}_1$ (red link) with the arrow pointing away from the link (labelled (ii)). The five cells of the box again relate to the five links of the network. This shows that if this link attempts to co-transmit with $\mathscr{L}_3$ (blue), $\mathscr{L}_3$ will fail. Further, consider the equivalent key attached to $\mathscr{L}_5$ (labelled (iii)). Here it is indicated that if this link attempts to co-transmit with $\mathscr{L}_1$, $\mathscr{L}_2$, $\mathscr{L}_3$ or $\mathscr{L}_4$ they will fail. Part (b) of each of the Figs. 7.2-7.9 indicates the maximal sets of co-transmitting links according to the theory developed in Sec. 7.2.

By considering the physical constraints, $\mathbf{E}$ and $\mathbf{F}$ matrices, defining collisions between links and links knowledge of the network respectively, can be generated, as can the symmetrised conflict matrix $\tilde{\mathbf{E}}$ (see Figs. 7.2-7.9 parts (d), (e) and (f)).

For each variation of the Braess network shown the capacity region is derived. The polytopes display feasible combinations of route flows (y_1, y_2, y_3) for the demand space $[0,1] \times [0,1] \times [0,1]$ (see Figs. 7.2-7.9 parts (c)). The volume of the capacity region $\tilde{V}$ is calculated in each case.

In turn Figs. 7.2-7.9 consider the following idealised interference scenarios:

1. Basic — see Fig. 7.2: nodes are in close physical proximity and there is no sort of interference management. It follows that every link is in conflict with every other link.

2. TP ($\mathscr{L}_2$,$\mathscr{L}_4$) — see Fig. 7.3: due to asymmetry of this topology some nodes may tune their TP (see Chapter 6 and [56, 142]) to improve spatial re-use, hence enabling some links to co-transmit: here, $\mathscr{L}_2$ and $\mathscr{L}_4$ together.

3. TP ($\mathscr{L}_1, \mathscr{L}_5$) — see Fig. 7.4: due to asymmetry of this topology (in the opposite direction to previously) some nodes may tune their TP to improve spatial re-use, hence enabling some links to co-transmit: here, $\mathscr{L}_1$ and $\mathscr{L}_5$ together.

4. Beamforming — see Fig. 7.5: transmitters use directional diversity techniques (see Chapter 5 and [133, 208, 260]) to minimise their interference on the rest of the network, hence enabling some links to co-transmit: here, $\mathscr{L}_1$ and $\mathscr{L}_5$ together, and $\mathscr{L}_2$ and $\mathscr{L}_4$ together.

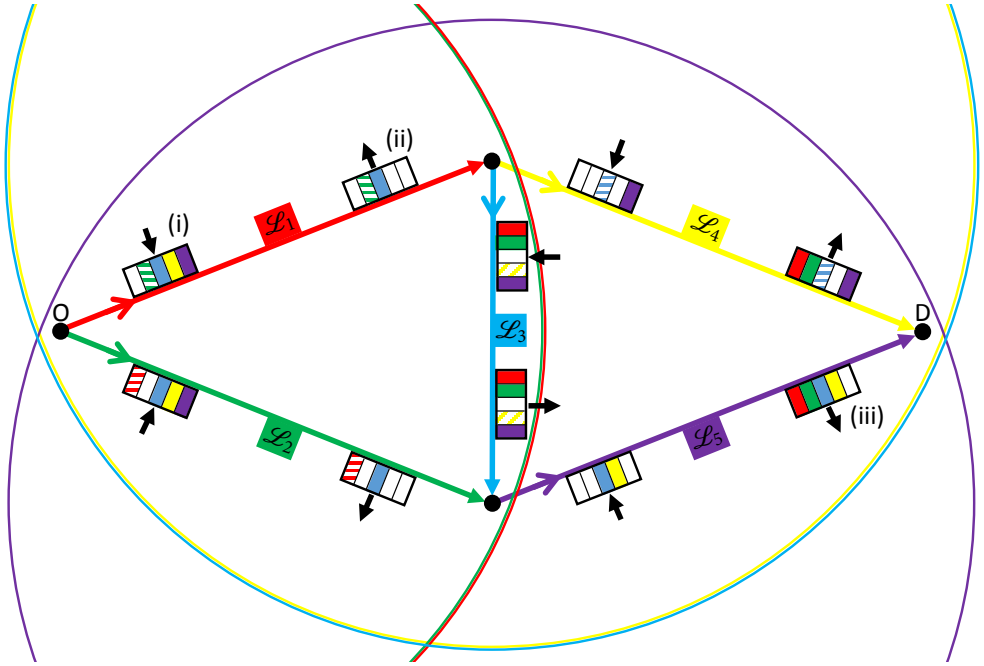
5. Full-duplex — see Fig. 7.6: nodes use uni-directional full-duplex (see Chapter 5 and [133, 152]) hence enabling some links to co-transmit: here, $\mathscr{L}_1$ and $\mathscr{L}_4$ together and $\mathscr{L}_2$ and $\mathscr{L}_5$ together.

6. Full-duplex with TP ($\mathscr{L}_2, \mathscr{L}_4$) — see Fig. 7.7: nodes use uni-directional full-duplex together with TP (combining ideas from Chapters 5 and 6 and [56, 133, 142, 152]) hence enabling some links to co-transmit: here, $\mathscr{L}_2$ and $\mathscr{L}_4$ together, $\mathscr{L}_1$ and $\mathscr{L}_4$ together and $\mathscr{L}_2$ and $\mathscr{L}_5$ together.

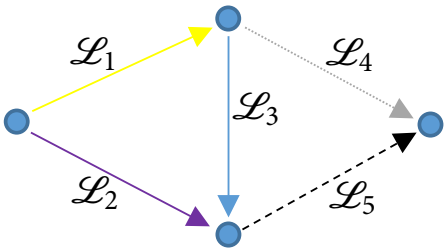
7. Full-duplex with TP ($\mathscr{L}_1, \mathscr{L}_5$) — see Fig. 7.8: nodes use uni-directional full-duplex together with TP hence enabling some links to co-transmit: here, $\mathscr{L}_1$ and $\mathscr{L}_5$ together, $\mathscr{L}_1$ and $\mathscr{L}_4$ together, $\mathscr{L}_2$ and $\mathscr{L}_5$ together and $\mathscr{L}_1$ and $\mathscr{L}_3$ together.

8. Full-duplex with beamforming — see Fig. 7.9: nodes use uni-directional full-duplex combined with beamforming (see Chapter 5 and [133, 134, 152]) hence enabling some links to co-transmit: here, $\mathscr{L}_1$ and $\mathscr{L}_3$ together, $\mathscr{L}_1$ and $\mathscr{L}_4$ together, $\mathscr{L}_2$ and $\mathscr{L}_5$ together, $\mathscr{L}_3$ and $\mathscr{L}_5$ together, $\mathscr{L}_1$ and $\mathscr{L}_5$ together, and $\mathscr{L}_2$ and $\mathscr{L}_4$ together.

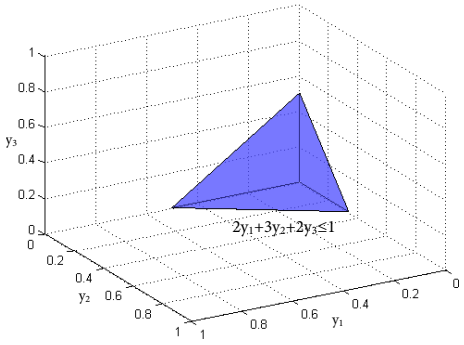
These eight settings form a progression where as the number of possible co-transmitting links increases, as does the capacity region. Setting 1. is the worst possible situation in which no co-transmissions are possible, and Setting 8. is the best possible situation, in which all co-transmissions are possible, except those that either originate at the same node or terminate at the same node.



(a)



(b)



(c)

$$\begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{pmatrix}$$

(d) $\mathbf{E}$

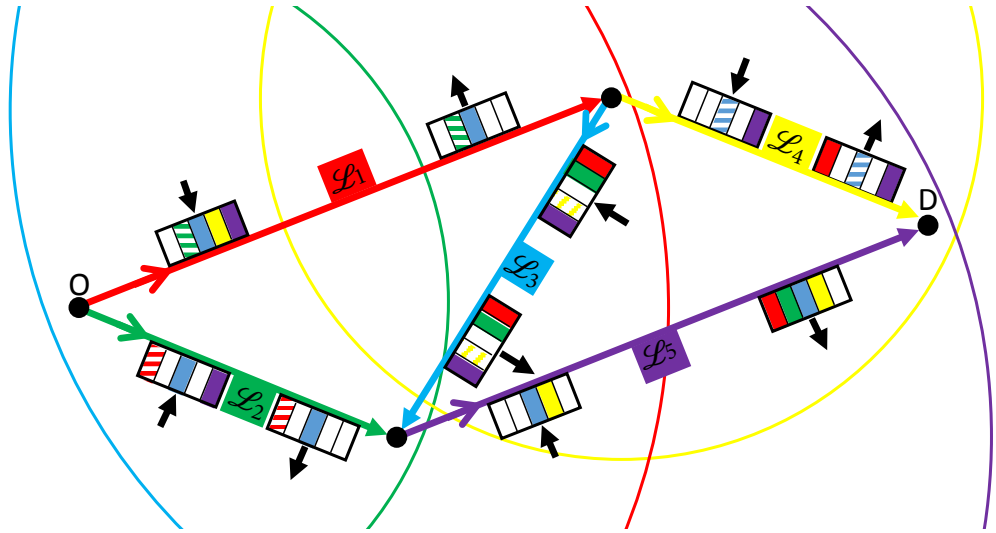
$$\begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 \end{pmatrix}$$

(e) $\tilde{\mathbf{E}}$

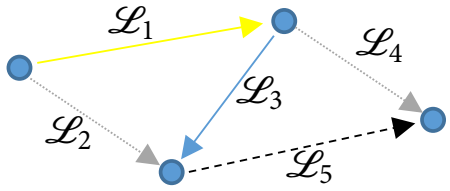
$$\begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 \end{pmatrix}$$

(f) $\mathbf{F}$

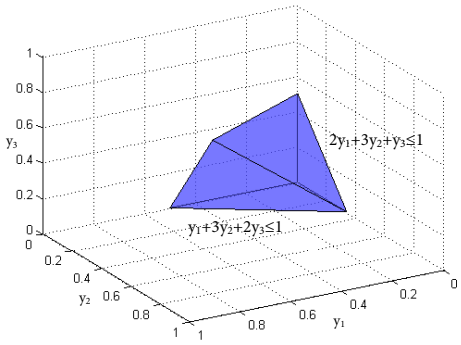
FIGURE 7.2. (a) Basic multi-hop network. Circles indicate the audible range of the corresponding colour link. (b) Sets of co-transmitting links. Links shown in the same colour / style can co-transmit. In this case, no pairs of links can co-transmit. (c) Capacity region with volume $\tilde{V} = 1/72$. (d) $\mathbf{E}$. (e) $\tilde{\mathbf{E}}$. (f) $\mathbf{F}$.



(a)



(b)



(c)

$$\begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{pmatrix}$$

(d) $\mathbf{E}$

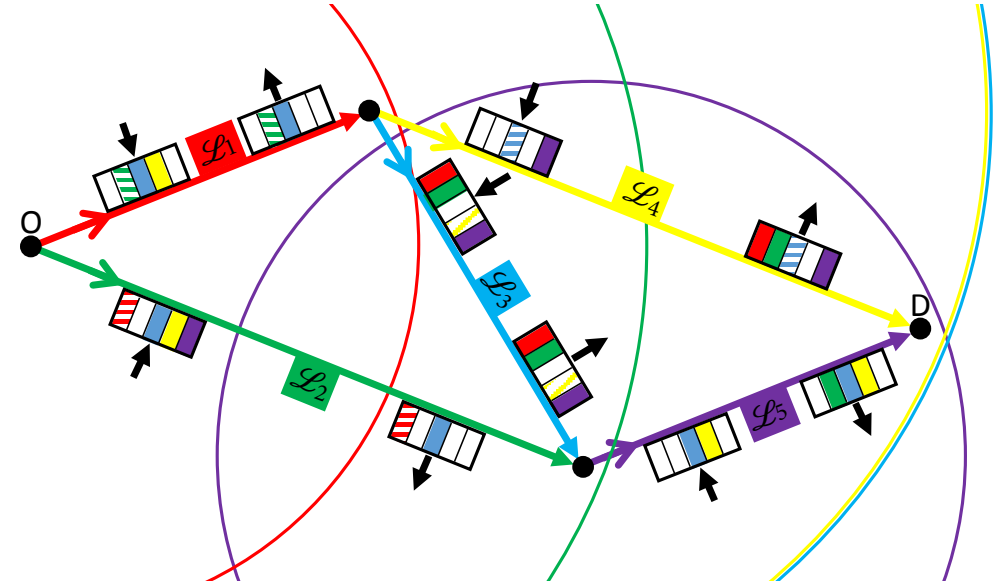
$$\begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 \end{pmatrix}$$

(e) $\tilde{\mathbf{E}}$

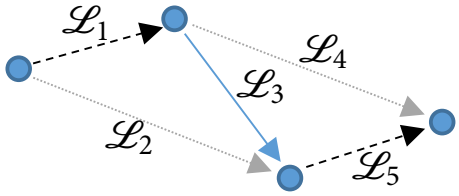
$$\begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 \end{pmatrix}$$

(f) $\mathbf{F}$

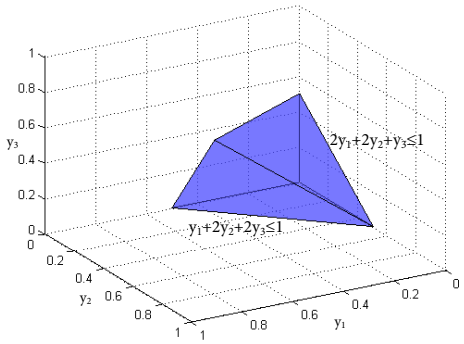
FIGURE 7.3. (a) TP $(\mathscr{L}_2, \mathscr{L}_4)$ multi-hop network. The northern and southern junction nodes are re-positioned such that their distances from the OD nodes differ. The link structure and possible routes through the network remain the same as previously (see Fig 7.2). Now the minimum power required for O to transmit to the southern junction node is less than that required to transmit to the northern junction. Similarly, the northern junction node transmitting to D requires less power than the southern junction does. Circles indicate the audible range of the corresponding colour link, illustrating how here $\mathscr{L}_2$ and $\mathscr{L}_4$ can reduce TP such that they may co-transmit. (b) Sets of co-transmitting links. Links shown in the same colour / style can co-transmit. (c) Capacity region with volume $\tilde{V} = 1/50$. (d) $\mathbf{E}$. (e) $\tilde{\mathbf{E}}$. (f) $\mathbf{F}$.



(a)



(b)



(c)

$$\begin{pmatrix} 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{pmatrix}$$

(d) $\mathbf{E}$

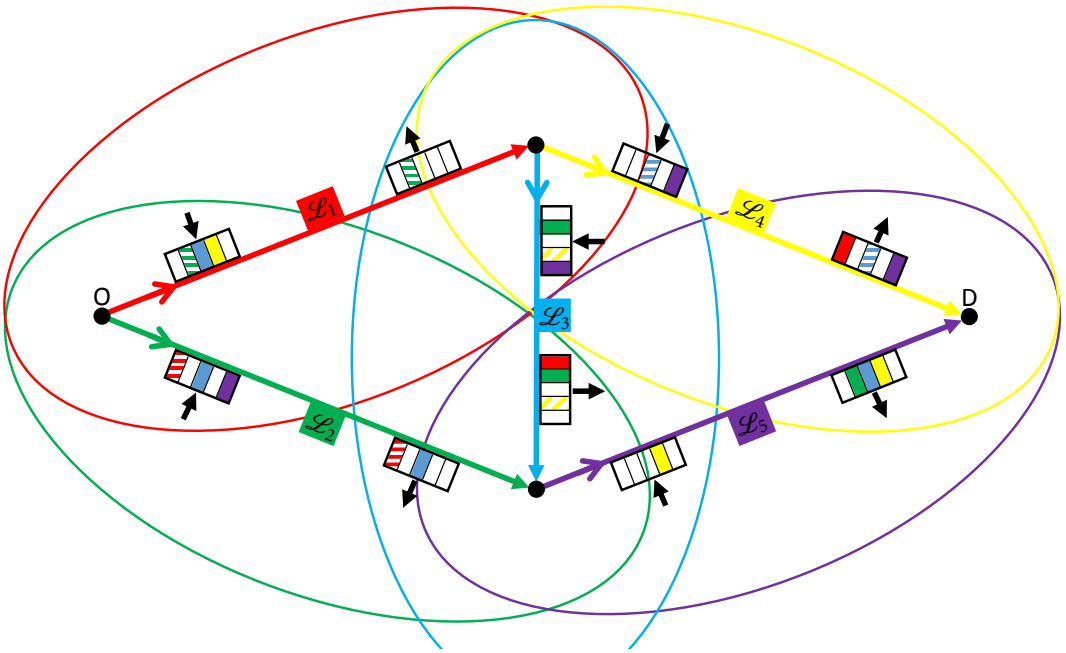
$$\begin{pmatrix} 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 \end{pmatrix}$$

(e) $\tilde{\mathbf{E}}$

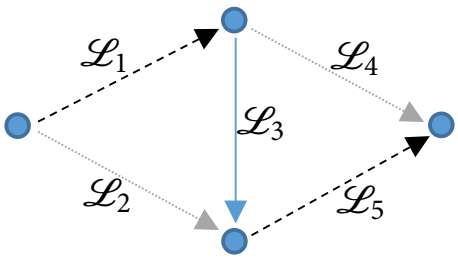
$$\begin{pmatrix} 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 \end{pmatrix}$$

(f) $\mathbf{F}$

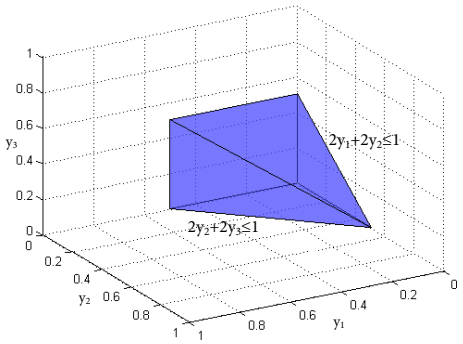
FIGURE 7.4. (a) TP adapt $(\mathscr{L}_1, \mathscr{L}_5)$ multi-hop network. The northern and southern junction nodes are re-positioned such that their distances from the OD nodes differ, presenting an alternative asymmetric variation of the previous network (see Fig. 7.3). Circles indicate the audible range of the corresponding colour link, illustrating how here $\mathscr{L}_1$ and $\mathscr{L}_5$ can reduce TP such that they may co-transmit. (b) Sets of co-transmitting links. Links shown in the same colour / style can co-transmit. (c) Capacity region with volume $\tilde{V} = 1/34$. (d) $\mathbf{E}$. (e) $\tilde{\mathbf{E}}$. (f) $\mathbf{F}$.



(a)



(b)



(c)

$$\begin{pmatrix} 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

(d) $\mathbf{E}$

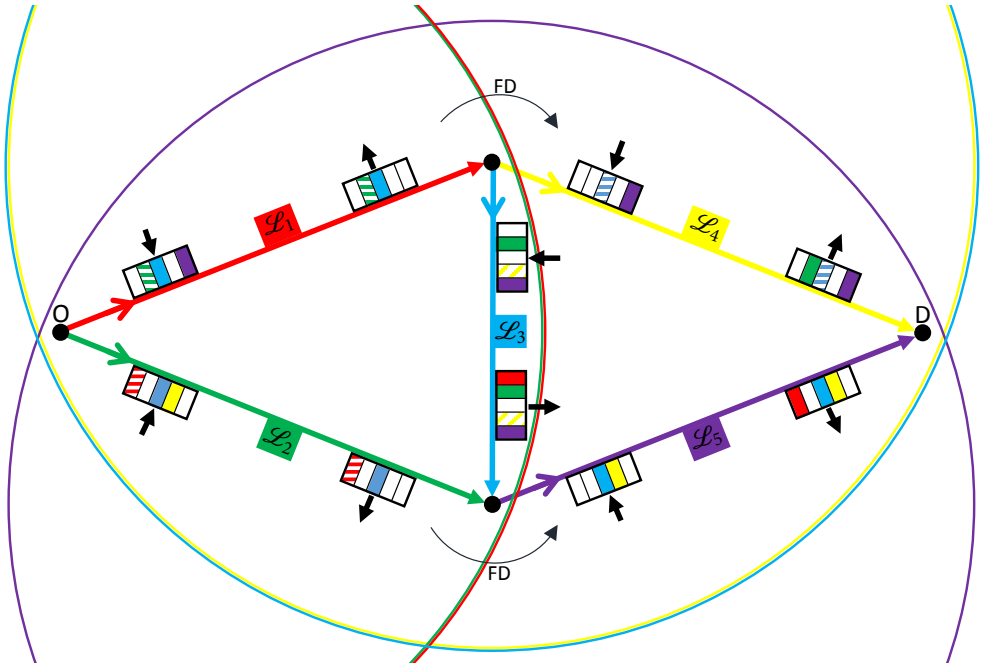
$$\begin{pmatrix} 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 \end{pmatrix}$$

(e) $\tilde{\mathbf{E}}$

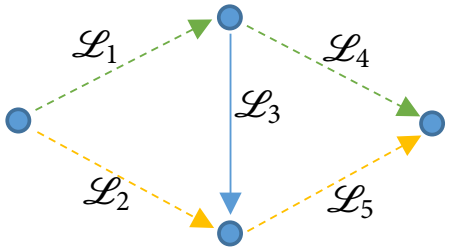
$$\begin{pmatrix} 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 \end{pmatrix}$$

(f) $\mathbf{F}$

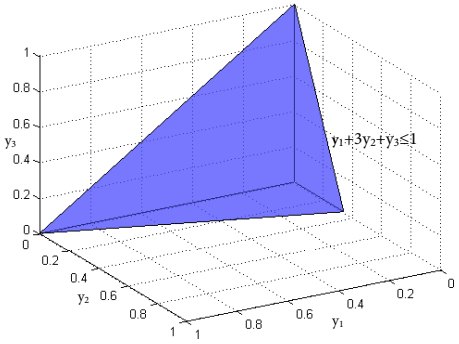
FIGURE 7.5. (a) Beamforming multi-hop network. Ovals indicate the directed audible range of the corresponding colour link, illustrating how beamforming changes the audible area of a transmission such that co-transmissions may occur for $\mathscr{L}_1$, $\mathscr{L}_5$ and $\mathscr{L}_2$, $\mathscr{L}_4$. (b) Sets of co-transmitting links. Links shown in the same colour / style can co-transmit. (c) Capacity region with volume $\tilde{V} = 1/24$. (d) $\mathbf{E}$. (e) $\tilde{\mathbf{E}}$. (f) $\mathbf{F}$.



(a)



(b)



(c)

$$\begin{pmatrix} 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{pmatrix}$$

(d) $\mathbf{E}$

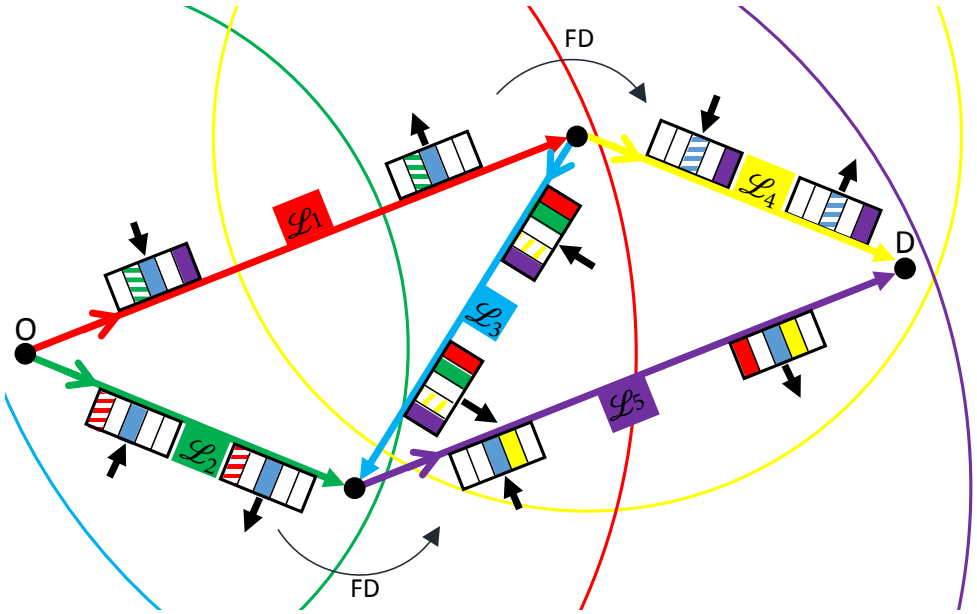
$$\begin{pmatrix} 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 \end{pmatrix}$$

(e) $\tilde{\mathbf{E}}$

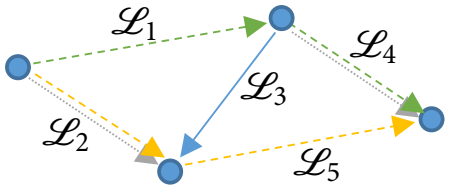
$$\begin{pmatrix} 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 \end{pmatrix}$$

(f) $\mathbf{F}$

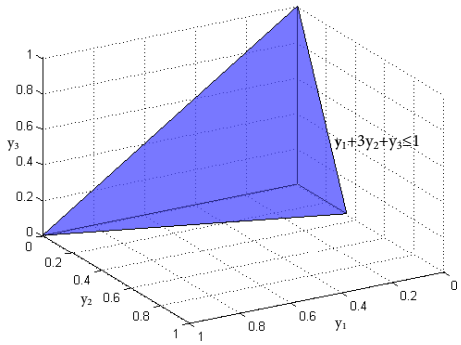
FIGURE 7.6. (a) Full-duplex multi-hop network. Additional arrows indicate how $\mathscr{L}_1$, $\mathscr{L}_4$ and $\mathscr{L}_2$, $\mathscr{L}_5$ can use unidirectional full-duplex, implemented at the northern and southern junctions, to co-transmit. (b) Sets of co-transmitting links. Links shown in the same colour / style can co-transmit. (c) Capacity region with volume $\tilde{V} = 1/18$. (d) $\mathbf{E}$. (e) $\tilde{\mathbf{E}}$. (f) $\mathbf{F}$.



(a)



(b)



(c)

$$\begin{pmatrix} 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{pmatrix}$$

(d) $\mathbf{E}$

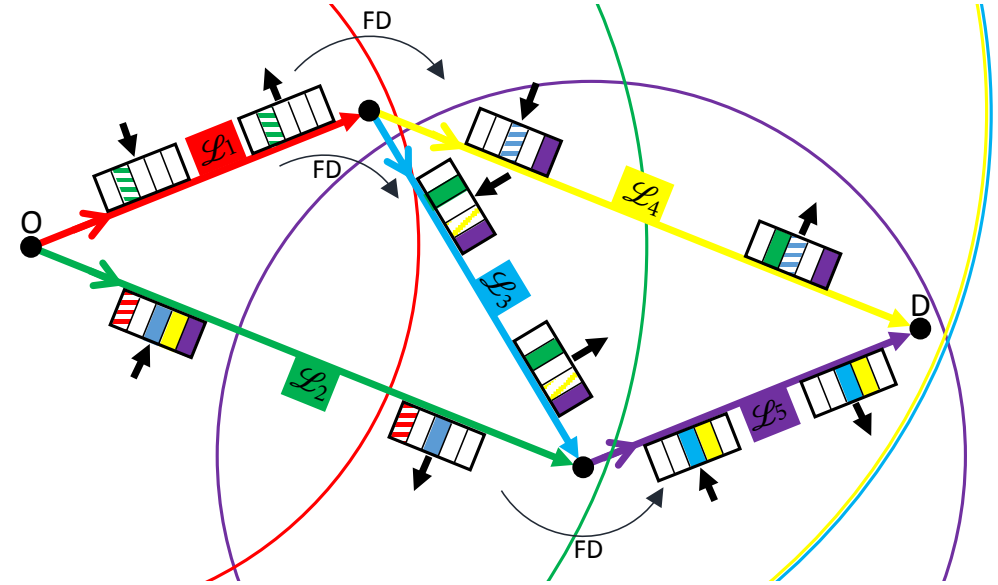
$$\begin{pmatrix} 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 \end{pmatrix}$$

(e) $\tilde{\mathbf{E}}$

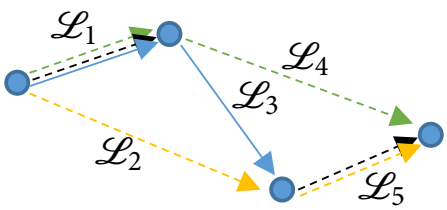
$$\begin{pmatrix} 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 \end{pmatrix}$$

(f) $\mathbf{F}$

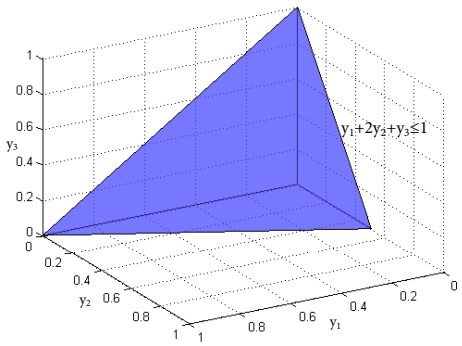
FIGURE 7.7. (a) TP $(\mathscr{L}_2, \mathscr{L}_4)$ (see Fig. 7.3) with full-duplex (see Fig. 7.6) multi-hop network. Circles indicate the audible range of the corresponding colour transmission, illustrating how here $\mathscr{L}_2$ and $\mathscr{L}_4$ can reduce TP such that they may co-transmit. Additional arrows indicate how $\mathscr{L}_1$, $\mathscr{L}_4$ and $\mathscr{L}_2$, $\mathscr{L}_5$ can use unidirectional full-duplex, implemented at the northern and southern junctions to co-transmit. (b) Sets of co-transmitting links. Links shown in the same colour / style can co-transmit. (c) Capacity region with volume $\tilde{V} = 1/18$. (d) $\mathbf{E}$. (e) $\tilde{\mathbf{E}}$. (f) $\mathbf{F}$. This is the first network in the progression in which individual links (here $\mathscr{L}_2$ and $\mathscr{L}_4$) appear in multiple co-transmitting sets.



(a)



(b)



(c)

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{pmatrix}$$

(d) $\mathbf{E}$

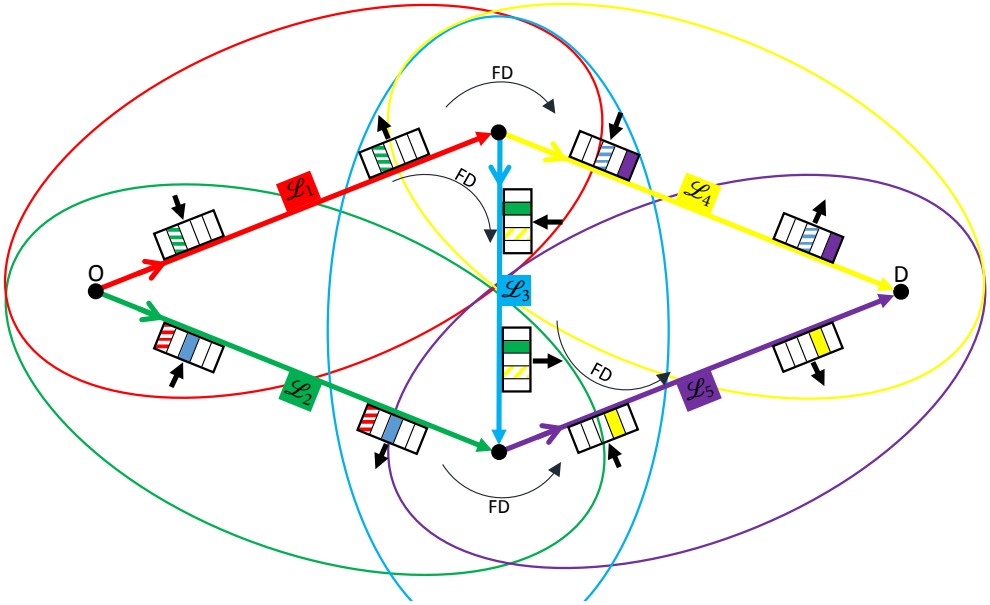
$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{pmatrix}$$

(e) $\tilde{\mathbf{E}}$

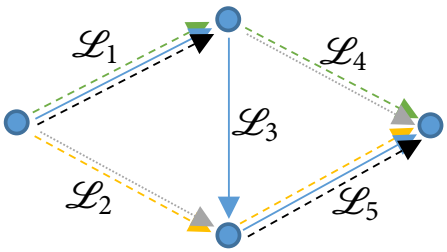
$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{pmatrix}$$

(f) $\mathbf{F}$

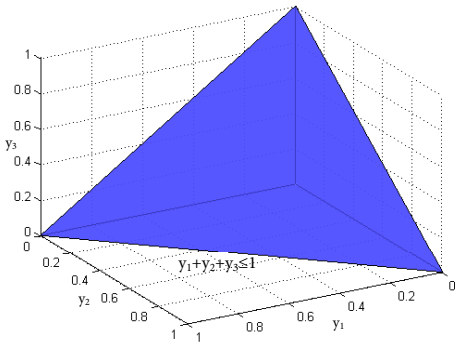
FIGURE 7.8. (a) TP $(\mathscr{L}_1, \mathscr{L}_5)$ (see Fig. 7.4) with full-duplex (see Fig. 7.6) multi-hop network. Circles indicate the audible range of the corresponding colour transmission, illustrating how here $\mathscr{L}_1$ and $\mathscr{L}_5$ can reduce TP such that they may co-transmit. Additional arrows indicate how $\mathscr{L}_1$, $\mathscr{L}_4$ and $\mathscr{L}_2$, $\mathscr{L}_5$ and further $\mathscr{L}_1$, $\mathscr{L}_3$ can use unidirectional full-duplex, implemented at the northern and southern junctions to co-transmit. (b) Sets of co-transmitting links. Links shown in the same colour / style can co-transmit. (c) Capacity region with volume $\tilde{V} = 1/12$. (d) $\mathbf{E}$. (e) $\tilde{\mathbf{E}}$. (f) $\mathbf{F}$.



(a)



(b)



(c)

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

(d) $\mathbf{E}$

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

(e) $\tilde{\mathbf{E}}$

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

(f) $\mathbf{F}$

FIGURE 7.9. (a) Full-duplex (see Fig 7.6) with beamforming (see Fig. 7.5) multi-hop network — the best possible situation in reducing interference. Ovals indicate the directed audible range of the corresponding colour transmission, illustrating how beamforming changes the audible area of a transmission such that co-transmissions may occur for $\mathscr{L}_1$, $\mathscr{L}_5$ and $\mathscr{L}_2$, $\mathscr{L}_4$. Additional arrows indicate how $\mathscr{L}_1$, $\mathscr{L}_4$ and $\mathscr{L}_2$, $\mathscr{L}_5$ and further $\mathscr{L}_1$, $\mathscr{L}_3$, $\mathscr{L}_5$ can use unidirectional full-duplex, implemented at the northern and southern junctions to co-transmit. (b) Sets of co-transmitting links. Links shown in the same colour / style can co-transmit. (c) Capacity region with volume $\tilde{V} = 1/6 = 1/3!$ which is the theoretical maximum for this network. (d) $\mathbf{E}$. (e) $\tilde{\mathbf{E}}$. (f) $\mathbf{F}$.

For the purposes of illustrating the detailed mathematical development outlined in Sec. 7.2, consider the beamforming scenario (Fig. 7.5). The symmetric conflict matrix takes the form

$$\tilde{\mathbf{E}} = \begin{bmatrix} 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 \end{bmatrix}, \tag{7.8}$$

where the zero off-diagonal entries in positions $(1,5)$ and $(2,4)$ (and symmetric counterparts) indicate pairs that can co-transmit. It is hence possible to then obtain three maximal compatible sets of links in the form

$$\mathbf{Z} = \left[\begin{array}{ccccc} 0 & 0 & 1 & 0 & 0 \\ \hline 0 & 1 & 0 & 1 & 0 \\ \hline 1 & 0 & 0 & 0 & 1 \end{array}\right], \tag{7.9}$$

which in this scenario is trivial, as there are no three-way compatible interactions. These are depicted by Fig. 7.5(b). By the substitution search procedure, the method then finds five spanning constraints on link flows in the form

$$\mathbf{C} = \left[\begin{array}{ccccc} 0 & 0 & 1 & 1 & 1 \\ \hline 0 & 1 & 1 & 0 & 1 \\ \hline 1 & 0 & 1 & 1 & 0 \\ \hline 1 & 1 & 1 & 0 & 0 \end{array}\right]. \tag{7.10}$$

For example, the first row of $\mathbf{C}$ states that the third, fourth and fifth columns of $\mathbf{Z}$ add up to (a vector less than or equal to) $\mathbf{1}_3$.

The route-link incidence matrix is

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \end{bmatrix}, \tag{7.11}$$

yielding four constraints on three route flows in the form of (7.7), where

$$\mathbf{C}\mathbf{A}^{\mathrm{T}} = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & 2 \\ 2 & 2 & 0 \\ 1 & 2 & 1 \end{bmatrix}. \tag{7.12}$$

However, the second and third constraints imply the first and fourth (identical) constraints. Hence it is possible to employ the reduced constraint matrix

$$\check{\mathbf{C}} = \begin{bmatrix} 0 & 2 & 2 \\ 2 & 2 & 0 \end{bmatrix}, \tag{7.13}$$

which defines the feasible polytope of route flows. See Fig. 7.5(c). If the objective is to maximise the OD flow $y_1 + y_2 + y_3$, the unique solution is $y_1 = y_3 = 1/2$ and $y_2 = 0$; i.e., traffic should be shared equally between the northern and southern routes, with the cross-town route left unused.

The introduction of beamforming shows a gain in capacity from the basic interference scenario (Fig. 7.2) which, in contrast yields a full $\mathbf{E}$ matrix and $\mathbf{Z}$ is a five-by-five identity matrix because no links may co-transmit. Subsequently the method obtains $\mathbf{C} = (1,1,1,1,1)$ (so the sum of link flows is bounded by one). When re-cast in route variables, this yields $2y_1 + 3y_2 + 2y_3 \leq 1$. Thus OD flow achieves its maximum of 1/2 for *any* combination of route flows with $y_1 + y_2 = 1/2$ and $y_3 = 0$.

The method illustrated here, and explained in detail in Sec. 7.2, could be applied to any of the described scenarios, or any other network, for which $\tilde{\mathbf{E}}$ can be defined (see Figs. 7.2-7.9 parts (e)). Of the eight described scenarios, the full-duplex with beamforming scenario (Fig 7.9) has the best (least) interference characteristics in which (7.8) is modified by further setting elements (1,3), (2,5) and (3,5) (and symmetric counterparts) equal to zero. This yields

$$\mathbf{Z} = \left[\begin{array}{ccccc} 0 & 1 & 0 & 0 & 1 \\ \hline 0 & 1 & 0 & 1 & 0 \\ \hline 1 & 0 & 0 & 1 & 0 \\ \hline 1 & 0 & 1 & 0 & 1 \end{array}\right] \quad \text{and} \quad \mathbf{C} = \left[\begin{array}{ccccc} 0 & 0 & 0 & 1 & 1 \\ \hline 0 & 0 & 1 & 1 & 0 \\ \hline 0 & 1 & 1 & 0 & 0 \\ \hline 1 & 1 & 0 & 0 & 0 \end{array}\right]. \tag{7.14}$$

Note in particular (row four of $\mathbf{Z}$) that links $\mathscr{L}_1$, $\mathscr{L}_3$, and $\mathscr{L}_5$ (constituting the cross-town route) may co-transmit. In route flows, the four constraints thus reduce to $y_1 + y_2 + y_3 \leq 1$, and maximum OD flow may be achieved by *any* route assignment.

7.4 Single OD Simulation Study

In a wireless network, a transmission will fail if the receiver is not able to hear the signal over interference and background noise. Consider the basic network topology shown in Fig. 7.2. If $\mathscr{L}_1$ and $\mathscr{L}_4$ have both sensed the channel clear and begin transmitting simultaneously, $\mathscr{L}_4$ will complete its transmission successfully but $\mathscr{L}_1$ will fail. This happens because the receiver of the transmissions on $\mathscr{L}_4$ cannot hear the transmissions occurring on $\mathscr{L}_1$, however the intended receiver of $\mathscr{L}_1$ is unable to receive due to the noise of the transmissions of $\mathscr{L}_4$ (the same is true of $\mathscr{L}_2$ and $\mathscr{L}_5$). This problem is reflected by asymmetry in the collision matrix $\mathbf{E}$ (Fig. 7.2(d)), however of course, is not captured by the the symmetrised $\tilde{\mathbf{E}}$, and so is not modelled by the theory developed in Sec. 7.2 and illustrated in Sec. 7.3.

It is important to note that when introducing new technologies to the network, in addition to enabling some links to transmit in parallel, further interference effects are changed. For example, implementing TP (as Fig. 7.4) enables $\mathscr{L}_1$ and $\mathscr{L}_5$ to transmit in parallel, but additionally impacts the symmetry of interference across the network. For example, if $\mathscr{L}_1$ and $\mathscr{L}_3$ begin transmitting simultaneously, whereas previously both would have mutually failed, now just $\mathscr{L}_1$ fails. This is because interference from $\mathscr{L}_1$ is no longer received at the destination of $\mathscr{L}_3$.

Similarly implementing beamforming (as Fig. 7.5) enables $\mathscr{L}_1$ and $\mathscr{L}_5$, as well as $\mathscr{L}_2$ and $\mathscr{L}_4$ to transmit in parallel. Now, as with TP, if $\mathscr{L}_1$ and $\mathscr{L}_3$ were to begin transmitting simultaneously, whereas previously both would have mutually failed, now just $\mathscr{L}_1$ fails as no interference from $\mathscr{L}_1$ is received at the destination of $\mathscr{L}_3$ (similarly $\mathscr{L}_3$ and $\mathscr{L}_5$). These changes introduce further asymmetry to the collision matrix.

Sec. 7.3 demonstrated an analytical method for computing the maximum achievable capacity region for a single OD example. The growth in this maximum achievable capacity offered by introducing new technologies is shown in Figs. 7.3 to 7.9. This section presents a more realistic simulation methodology to identify the capacity of a multi-hop CSMA/CA network for comparison with analytical results.

7.4.1 Simulating Multi-Hop Networks

The conceptual simulation modelling methodology presented in Chapter 4 is extended to multi-hop networks (for all simulations in this chapter $Countdown_{\max} = 2$). Traffic is prescribed to each link using the route-link incidence matrix (7.11), such that when demand is prescribed to a route, an equal demand is prescribed to each link in that route. This method provides a simple way of observing effects of competing route demands, without having to track individual packets from origin to destination.

Simulations are based on the network structure defined by the **E** and **F** matrices (see Figs. 7.2 to 7.9). For each network structure, two options are considered. The first supposes that links originating from the same node share a single queue. Thus, (for example) O is assumed to have a queue awaiting transmission, with Poisson arrival rate, consisting of a mix of packets to be sent via $\mathscr{L}_1$ and $\mathscr{L}_2$, the ratios of which depending on the demands on each of the three routes. Packets are then served in order of arrival, irrespective of their route (no active queue management). Hence at O, packets bound for the southern junction node may be blocked by packets bound for the northern junction node if $\mathscr{L}_1$ is prone to collisions, even if $\mathscr{L}_2$ is relatively collision-free. (Similarly the northern junction node maintains a single queue serving $\mathscr{L}_3$ and $\mathscr{L}_4$.)

The second option (queue management) supposes that each link maintains its own independent queue; thus (for example) O is assumed to have *two* queues of packets awaiting transmission, with Poisson arrival rates, for $\mathscr{L}_1$ and $\mathscr{L}_2$ respectively. Hence, if $\mathscr{L}_1$ is prone to collisions, packets bound for the northern junction may be blocked, but $\mathscr{L}_2$ can continue to serve traffic bound for the southern junction regardless. However, a node can only transmit along one link at any one time. Assuming a packet is queued for each link and there are no obstructions when the channel is sensed clear, O randomly selects to transmit on either $\mathscr{L}_1$ or $\mathscr{L}_2$ such that over a long simulation time, they each receive equal opportunity to transmit. (Similarly, the northern junction node now maintains two queues serving $\mathscr{L}_3$ and $\mathscr{L}_4$.)

For each of the cases shown in Figs. 7.2 to 7.9, both single and multiple queue options are considered, and a large ensemble of simulations are performed each with a different route demand

intensity vector (y_1, y_2, y_3) varying traffic applied to each route from 0 to 1. Here 1 indicates the maximum rate at which a route can carry data traffic when transmitting alone in an otherwise clear channel before a queue starts to build. Analogous to previous simulation studies in this thesis, each of the route demands y_1, y_2 and y_3 is varied independently from 0 to 1 in 40 equal increments, resulting in $41^3 = 68,921$ simulations. The duration of each simulation is set at the time needed to transmit $10,000$ packets without collisions.

7.4.2 Simulation Results and Analysis

Fig. 7.10 illustrates the simulation results, and the overall network structure depicts the development of $\tilde{\mathbf{E}}$: as one follows an arrow connecting the various technology scenarios to the right, interference is reduced. Precisely, the $\tilde{\mathbf{E}}$ matrix, defining the technology scenario with an arrows pointing into it, contains '0's in the same position as the scenario the arrow originates from as well as further additional '0's.

The simulation results are presented in the form of three-dimensional scatter plots. Figs. 7.10(a)-(g) show this for a sample of the technology cases contained within the polytopes derived analytically. To better understand the results, numerical measures are provided as follows: $|\tilde{E}|$ which denotes the number of '1's in one half of the symmetrical conflict matrix, and further $\tilde{V}$, V and $\ddot{V}$ denote the volume of the analytically derived polytope, the simulated results (considering a single queue on each node) and the simulated results with queue management (i.e., multiple queues on nodes that serve multiple links) respectively. The proportions of the polytope achieved by the simulations are then calculated, i.e., $V/\tilde{V}$ and $\ddot{V}/\tilde{V}$.

Simulation of the basic network topology showed route demands able to achieve just more than half of the analytically calculated maximum capacity. Jindal and Psounis [129] identified a similar ratio comparing optimal calculated throughput to that of the throughput achieved from their multi-hop network model. As only one link is capable of transmitting successfully at any one time in the basic topology, the introduction of queue management was unable to realise any gain.

Introducing new technologies to the simulation can remove some of the constraints on the networks, allowing pairs of links to transmit in parallel, and thus increases the size of the polytopes defining the maximum achievable network capacity. For example, the introduction of TP adaptation increases the size of the polytope from the basic case (Figs. 7.3 and 7.4). Simulation results however, for the case of a single queue on each node, show no gain in the capacity region when TP adaptation was implemented. Further, for beamforming (Fig. 7.5), only a very minor gain was observed. Only with the further addition of queue management did the simulation results show the ability to realise the benefit of the new technology and more significant growth in the volume of the capacity region. Figs. 7.10(b) and 7.10(c) illustrate the difference in capacity region volume, with and without queue management, for the beamforming case. A similar impact can be observed comparing Figs. 7.10(d) and 7.10(e), and further Figs. 7.10(f) and 7.10(g), which

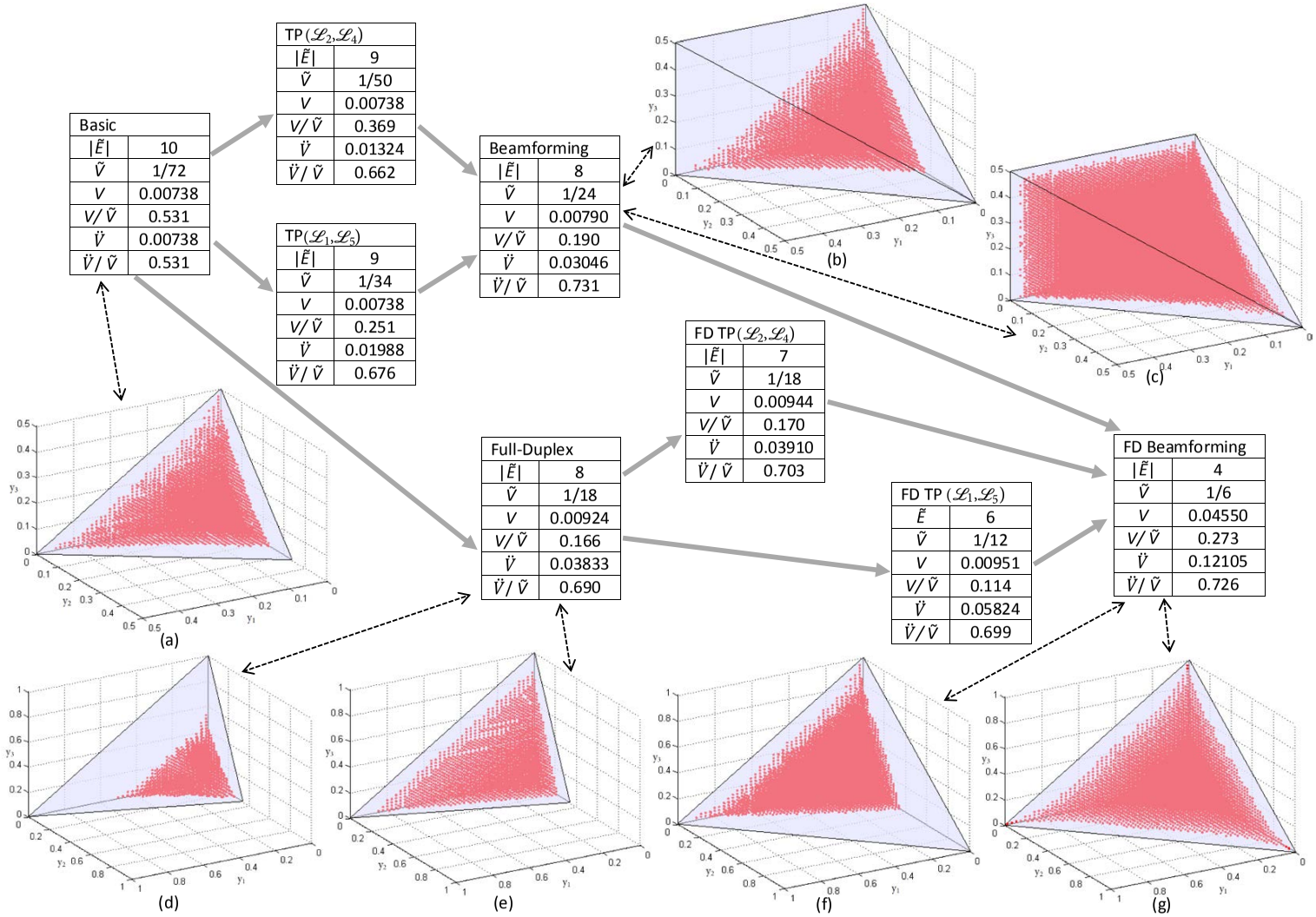


FIGURE 7.10. Single OD multi-hop network capacity region simulation results. $|\tilde{E}|$ specifies the number of '1's in one half of the symmetrical conflict matrix. $\tilde{V}$ is the volume of the polytope. V is the volume of the simulated results when each node maintains a single queue. $\ddot{V}$ the volume of the simulated results with queue management (i.e., multiple queues on nodes that serve multiple links). As you follow an arrow between tables to the right, the scenario's matrix $\tilde{\mathbf{E}}$ is a development of previous cases, i.e., $\tilde{\mathbf{E}}$ contains '0's in the same position and further additional '0's, hence $|\tilde{E}|$ is less. Simulated capacity regions are shown contained within the polytope for basic, beamforming, full-duplex (FD) and full-duplex with beamforming scenarios. Red markers indicate simulated combinations of route demands measured inside the capacity region. Sub figures (a) (b) (d) (f) show single mixed traffic queue cases, whereas (c) (e) (g) show queue management cases.

show the simulated results for both full-duplex (see Fig. 7.6) and full-duplex with beamforming (see Fig. 7.9) scenarios. For each technology, or combination of technologies, introduced to the basic multi-hop example, the system fails to realise significant gain without queue management.

To understand these results in more detail, consider the beamforming case. Link pairs $\mathscr{L}_1$, $\mathscr{L}_5$ and $\mathscr{L}_2$, $\mathscr{L}_4$ can transmit in parallel (see Fig. 7.5). When just a single queue is maintained on each node, the simulation results show that the introduction of beamforming achieves only slight growth in the capacity region relative to the basic topology. This is due to a lack of coordination between nodes, which all sense the channel independently and attempt to transmit along the link dictated by the packet at the front of their queue. For example, if O is transmitting to the southern junction node via $\mathscr{L}_2$, the northern junction node has the capability to transmit along $\mathscr{L}_4$ simultaneously. However, if the packet at the front of the northern junction's queue is for $\mathscr{L}_3$, it must wait for the transmission on $\mathscr{L}_2$ to finish then contest the channel again before it can attempt to transmit, even if there are other packets further back in its queue bound for D along $\mathscr{L}_4$. This lack of coordination between nodes means the system fails to realise any significant benefit from beamforming. When queue management is enabled this is no longer an issue. Rather, any time O is transmitting to the southern junction via $\mathscr{L}_2$, the northern junction can transmit any packets destined for D along $\mathscr{L}_4$ by prioritising them to the front of the queue. Now the system can benefit from links transmitting in parallel. The compounded benefit of enabling simultaneous transmissions is that it further reduces contention at other times, hence reducing the number of collisions. In the beamforming network, the only occasions nodes contest the channel for individual channel access are when the northern junction wishes to transmit to the southern junction via $\mathscr{L}_3$. Further, the asymmetry of $\mathbf{E}$ means that in some cases, when two incompatible links do begin transmitting simultaneously, only one link will fail due to interference at the receiver. Therefore in this case the time is not completely wasted as one link can still successfully complete its transmission.

The introduced technologies are shown to provide significant gains to the system when queue management is enabled. Figs. 7.10(b) and 7.10(c) illustrate the difference in capacity region volume for the single mixed traffic queue and queue management simulations for the beamforming case. The capacity region volume more than doubles when queue management is introduced and achieves a much greater proportion of the analytically derived maximum polytope. Queue management is shown to have a similar effect, facilitating other technologies investigated to increase the system capacity and achieve a much greater proportion of their analytically derived maximum polytope.

When queue management is implemented, simulation results can produce capacity regions which fill a significant portion of the analytically derived polytope. Therefore, although it must be accepted that in reality the analytically derived maximum will never quite be achieved, the method presented in Sec. 7.2 seems to be a useful tool for exploring route constraints in multi-hop wireless networks and hence the chapter proceeds using this analytical approach to further

investigate routing for multiple OD pairs.

7.5 Designed Networks With Two OD Pairs

The investigation now considers wireless multi-hop networks with two OD pairs. The challenge to overcome is that the pairs' data flows have some nodes in common and thus contend with each other. In some cases, extra capacity can be gained if each OD pair distributes its traffic over several routes. The approach in this section is to illustrate principles by which networks can be designed to allow extra route choice and overcome contention. The next section (Sec. 7.6) considers whether these principles are achieved on randomly generated networks.

The remainder of the chapter supposes that all nodes have an idealised full-duplex and beamforming capability, as illustrated for the Braess network in Fig. 7.9. With beamforming technology alone, a pair of links is in conflict if they share a node. In contrast, with full-duplex and beamforming technology, a pair of links are in conflict only if either (a) they share a source node or (b) they share a sink node.

In the four-link network (Fig. 7.11(a)), the two OD pairs (OA,DA) and (OB,DB) are each connected by one route. When operating alone, either route could carry flow continuously with both links operating simultaneously in full-duplex. However, the achievable flow for both ODs together is limited by the shared relay node (i) which can only serve one input flow and one output flow at any time. Hence, the sum of the two OD flows cannot exceed the maximum achievable by one OD pair alone.

In Fig. 7.11(b), the pair (OA,DA) should maximise throughput by dividing its flow between the two shortest (two-link) routes via the relay nodes (ii). For illustration, suppose the (OA,DA) flow is maximal (one). The flow on each route is thus one half, and thus each of the relays is free half of the time to serve the flow for (OB,DB). In fact, (OB,DB) is able to achieve a flow of one half by coordinating with (OA,DA) and having its middle link active half of the time and its first and third links active the other half of the time. In contrast, if all the flow for (OA,DA) went via one route, (OB,DB) would be entirely shut out at one of the relays.

In Fig. 7.11(c), two 'splitting' nodes (iii), two 'joining' nodes (iv), and four relay nodes (v) are introduced so that each OD pair has two routes of five links each and two further routes of six links which remain unused. Maximal demand (one) may now be served for both OD pairs, by sharing each relay node's time accordingly. This requires each route to have its first, third and fifth links active half of the time and its second and fourth links active half of the time, coordinating with the other OD pair to avoid conflicts.

Fig. 7.11(d) displays the capacity regions for these three networks — that is, feasible combinations of OD flows $(d_{\mathrm{A}}, d_{\mathrm{B}})$ supposing that the OD pairs coordinate their route assignments. The boundaries are the Pareto fronts of GAME 1 in Sec. 7.2. Note the capacity region increases from a theoretical minimum to a theoretical maximum progressing through the three exemplars.

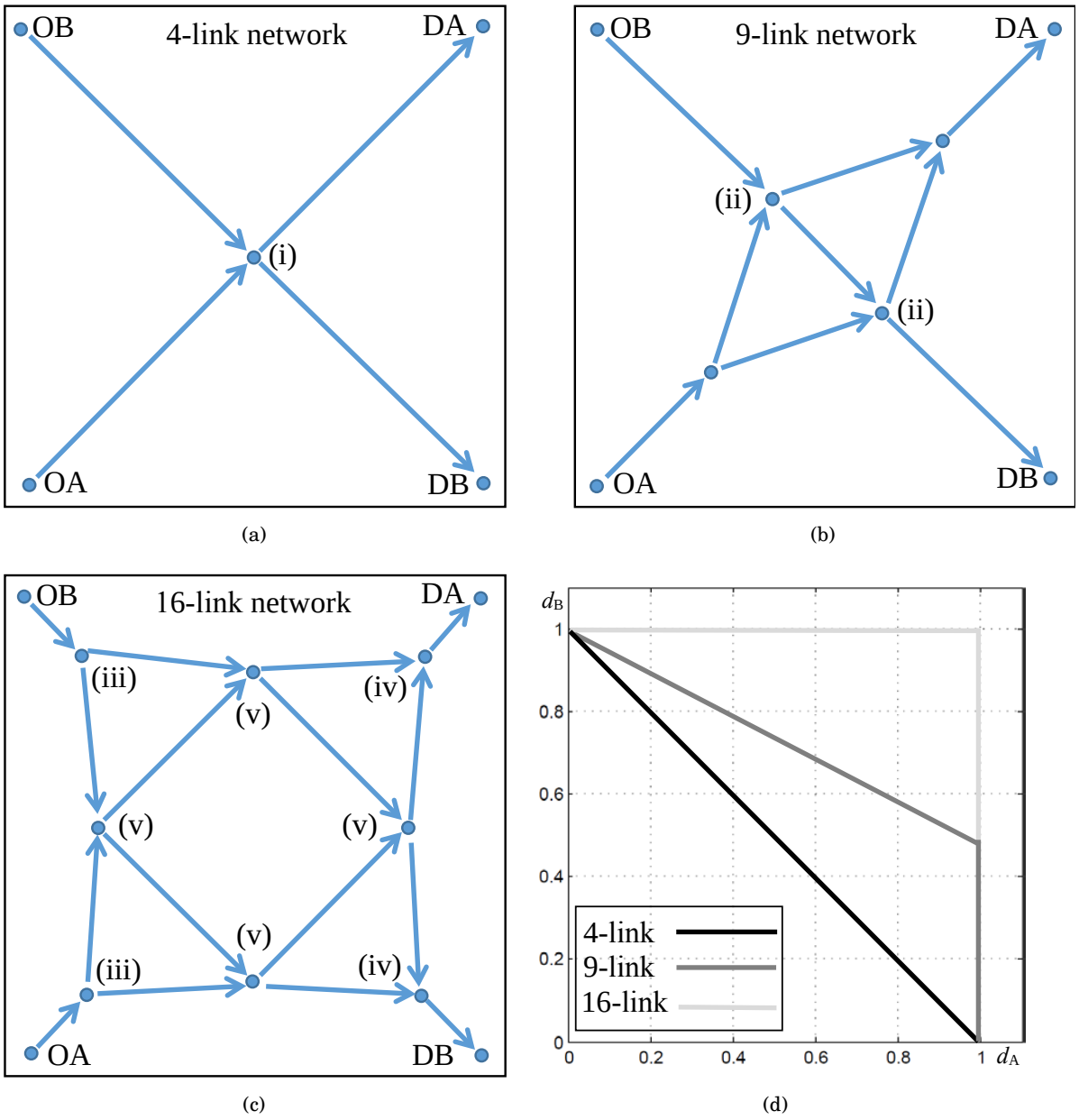


FIGURE 7.11. Exemplar networks connecting two OD pairs (OA,DA) and (OB,DB). (a) Simplest network with one shared relay node and one route for each OD pair. (b) Extra relay nodes are added. The pair (OA,DA) gains extra routes. (c) Eight interior nodes (4 shared) yielding multiple routes for both OD pairs. (d) Increase in capacity regions as relay nodes are added, assuming full-duplex and beamforming technology. Network (c) allows simultaneous full data rates on both OD pairs.

7.6 Random Networks With Two OD Pairs

It was shown how the introduction of additional relay nodes can enable an increase in capacity when two OD pairs are in contention across designed wireless multi-hop networks. The benefit is realised by splitting each OD flow across multiple routes. The aim now is to analyse how often this gain might be achieved in practice.

To this end, consider a similar set-up to Sec. 7.5, where two OD pairs (OA,DA) and (OB,DB) are positioned at diagonally opposite corners of the unit square $[0,1]\times[0,1]$. Then, position m relay nodes randomly in $[0.1,0.9]\times[0.1,0.9]$, which are joined to each other by Delaunay triangulation. This interior network is joined to the corners of the unit square by the closest relay nodes. To simplify matters, suppose each link in the network is unidirectional with the flow direction from left to right. For example, see Fig. 7.12(a).

The resulting network is a DAG for which all possible routes joining OA to DA and OB to DB are computed. Because the investigation is interested in the effect of route choice, networks which have only one or no routes for either OD pair are discarded. The process is repeated until 1,000 exemplars have been generated for each $m = 3,4,\ldots,8$. When $m = 3$, this involves generating about 7,700 networks, reducing to about 1,800 for $m = 8$. However, note that for small m, only a handful of topologically distinct cases result.

The computational process that follows is best illustrated by the example shown in Fig. 7.12(a), which has $m = 8$ relay nodes, $n = 21$ links, $r_{\mathrm{A}} = 14$ routes for (OA,DA) and $r_{\mathrm{B}} = 10$ routes for (OB,DB). Note that the process may generate redundant links — in this case, link $\mathscr{L}_6$ is not in any route from OA to DA or OB to DB. However, it is not removed from the analysis at this stage, because a single-hop flow upon it would impact on flows of interest via the relay nodes at its ends. A conflict matrix $\tilde{\mathbf{E}}$ is generated assuming all nodes have full-duplex and beamforming capability, and the analytical framework (Sec. 7.2) yields $p = 119$ maximal co-transmitting sets, $q = 15$ constraints on link flows, but only $\tilde{q} = 5$ reduced constraints on route flows.

Fig. 7.12(b) shows capacity region computations for this network — that is, feasible combinations of OD flows $(d_{\mathrm{A}}, d_{\mathrm{B}})$ which can be met by the OD pairs coordinating their route assignments cooperatively. The boundaries of the capacity regions are computed by playing the symmetric variants of GAME 1 for d_{A} (resp. d_{B}) running from 0 to 1 in increments of 0.01. The Pareto front computation is then trivial because the sloped sections of the boundary coincide in the symmetric game variants. Each boundary consists of straight line segments along which the route assignment may be fixed. Note that to achieve the boundary, at most $\tilde{q}+1 = 6$ routes are required, in accordance with theory (p.183), although sometimes optimality is achieved with even fewer routes. To simplify matters in the statistics that follow, the capacity region is summarised by the solution to GAME 2, that is the maximum feasible value of $d_{\mathrm{A}} + d_{\mathrm{B}}$ where $d_{\mathrm{A}} = d_{\mathrm{B}}$ as indicated on the figure.

Fig. 7.12(b) also shows how the capacity region shrinks as background single-hop traffic is introduced. Note that the number of used routes required for optimality also tends to decrease.

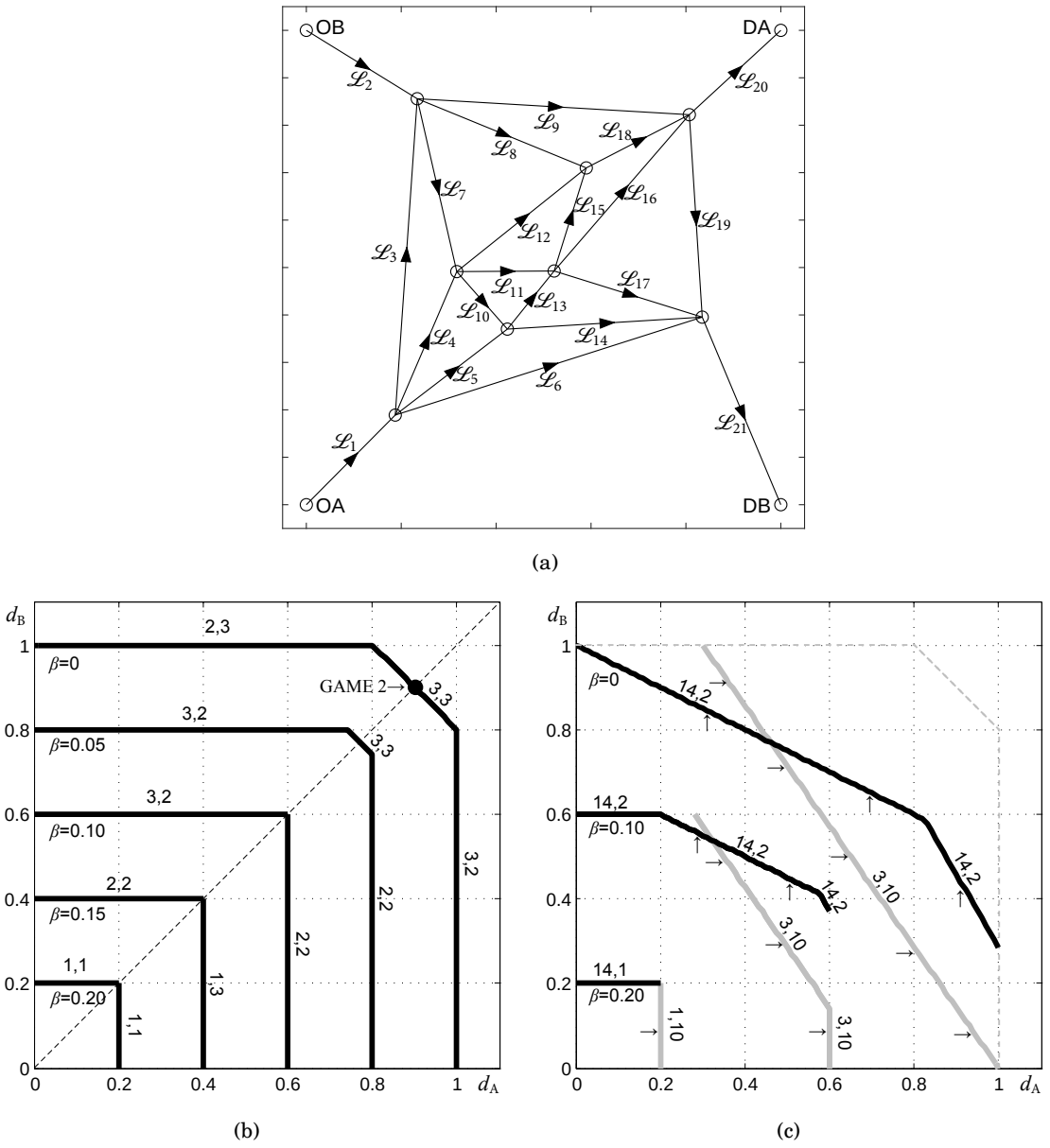


FIGURE 7.12. (a) Randomly generated network with $m = 8$ interior nodes and $n = 21$ links, resulting in $r_A = 14$ routes connecting (OA,DA) and $r_B = 10$ routes connecting (OB,DB). (b) Capacity regions when two OD pairs route cooperatively (GAME 1), as background traffic β is introduced. The boundaries are annotated with the minimum number of routes r_A^U, r_B^U that must be used for each OD pair in order to achieve the boundary. (c) Capacity regions when demand is fixed equally across all routes for one OD pair whilst maximising the sum of flows for the other OD pair (GAME 3).

Table 7.1: Network statistics for experimental ensembles of random wireless multi-hop networks. The parameter m describes the number of relay nodes. Other parameters are as defined in Sec. 7.2. Results reported to three significant figures; integer results are exact.

m	$\langle n \rangle$	$\langle r \rangle$	$p_{\min}$	$\langle p \rangle$	$p_{\max}$	$\langle q \rangle$	$\langle \tilde{q} \rangle$	P($\tilde{q}$=1)
3	7	4	8	8	8	4	1	1
4	9.51	5.76	10	12.2	20	6.71	1	1
5	11.9	7.26	9	18.9	44	8.68	1.57	0.715
6	14.5	9.98	12	32.5	100	10.5	2.24	0.514
7	17.1	13.4	21	59.9	216	12.3	2.85	0.379
8	19.8	17.8	28	115	365	14.0	3.33	0.300

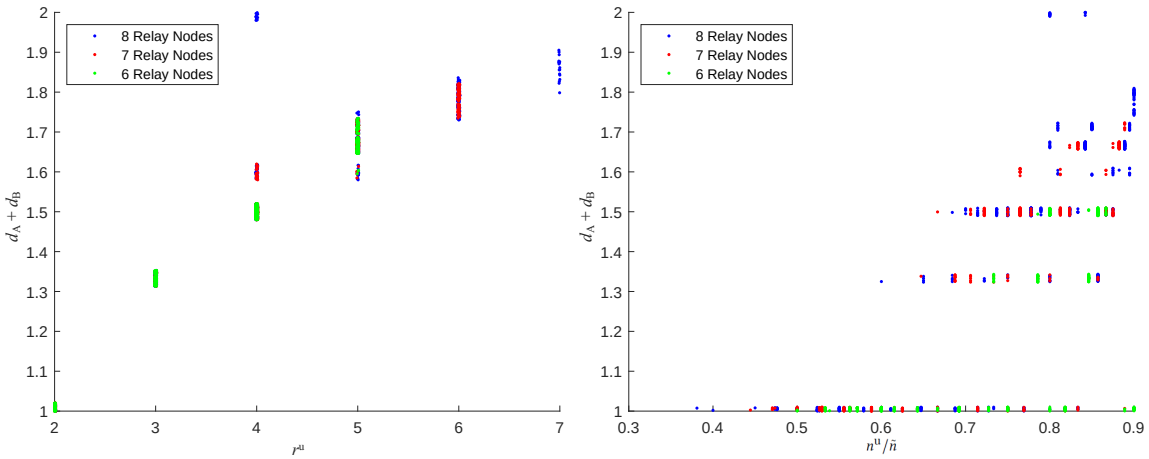
Table 7.2: Capacity statistics for experimental ensembles of random wireless multi-hop networks, characterised by maximising $d_{\mathrm{A}} + d_{\mathrm{B}}$ subject to $d_{\mathrm{A}} = d_{\mathrm{B}}$ (GAME 2). Here the superscript u refers to the number used in the optimal assignment and the subscript AB refers to the number shared in the optimal assignment. The parameter $\tilde{n} \leq n$ is the number of non-redundant links.

m	$\langle d_{\mathrm{A}} + d_{\mathrm{B}} \rangle$	$\langle r^{\mathrm{u}} \rangle$	P($r^{\mathrm{u}} = \tilde{q} + 1$)	$\langle m^{\mathrm{u}}_{\mathrm{AB}} \rangle$	$\langle n^{\mathrm{u}}_{\mathrm{AB}}/\tilde{n} \rangle$
3	1	2	1	3	0.286
4	1	2	1	3.19	0.231
5	1.13	2.47	0.906	3.34	0.175
6	1.25	3.03	0.794	3.65	0.134
7	1.37	3.59	0.745	4.01	0.109
8	1.44	3.93	0.651	4.28	0.0895

Fig. 7.12(c) shows results for the non-cooperative GAME 3, where one OD pair maximises its flow when the other presents to it a fixed assignment that distributes demand equally over routes. In contrast to GAME 1, the symmetric variants of the game can achieve different capacity bounds. Moreover, since the optimisations are over a strict subset of the variables used in GAME 1, where all route flows are free, the capacity bounds are typically lower.

Now consider statistical results for the 1,000 networks ensembles. Tab. 7.1 displays numbers of links n, possible routes r, co-transmitting sets p, and constraints (q in link flows and $\tilde{q}$ for the reduced route flow system) — i.e., parameters computed according to the theory developed in Sec. 7.2. Of course, the numbers of links and routes tend to increase with the number m of relay nodes. However (reflecting the known computational complexity), the number of co-transmitting sets grows very markedly, and moreover varies enormously from one example to the next, reflected in the observed range $[p_{\min}, p_{\max}]$. However, the constraint systems, particularly when expressed in route variables, are much smaller.

A notable feature is the proportion P($\tilde{q} = 1$) of networks with just one reduced constraint. This constraint is then always $\sum y_i = 1$. In such networks, the capacity region is $d_{\mathrm{A}} + d_{\mathrm{B}} \leq 1$, irrespective of the routing strategy employed. Encouragingly, the proportion of such networks decreases as the number of relay nodes increases, implying the potential for more frequent capacity gains from intelligent routing.



(a) (b)

FIGURE 7.13. Scatter plots for experimental ensembles of random wireless multi-hop networks. Results shown for $m = 6, 7, 8$ relay nodes only. GAME 2 capacity $d_A + d_B$ against (a) the number of routes r^u used in the optimal assignment and (b) the proportion $n^u/\tilde{n}$ of used links.

Tab. 7.2 considers the statistics of the optimal assignments on the 1,000 networks ensembles. Specifically, this solves GAME 2 by maximising $d_A + d_B$ (by which capacity is characterised) subject to $d_A = d_B$, and the numbers of used routes and shared relay nodes and links are studied.

Generally speaking, capacity increases with an increase in the number of relay nodes. This seems to be achieved by increasing the number of routes used in the optimal assignment. Moreover, in the optimal assignments, it seems $\langle n^u_{AB}/\tilde{n}\rangle \to 0$ as m increases, indicating that the OD flows approach a perfect separation at link level for larger networks. Of course, complete separation at node level is *not* possible, because the A and B flows must cross somewhere. Rather, it is observed that the number of shared used nodes $\langle m^u_{AB}\rangle \sim 4$, compatible with two routes for each OD pair crossing each other in the minimum number of places. All these results suggest larger random networks approach the optimal behaviours designed in Sec. 7.5.

Fig. 7.13(a) displays how improvements in capacity are distributed across the ensembles and emphasises the correlation with the number of used routes. Further, Fig. 7.13(b) shows that higher capacity exemplars tend to use a high proportion of their links in their optimal assignments, as might be expected. These plots exhibit a small number of exemplars for which the theoretical maximum $d_A + d_B = 2$ is attained. These are all for $m = 8$ relay nodes and correspond to two used routes for each OD pair — analogous to the designed example Fig. 7.11(c) — and as such, buck the overall statistical trend.

Note as previously discussed, at most $\tilde{q} + 1$ routes are required in the GAME 2 optimal assignment. However, this is an upper bound and (Tab. 7.2) larger networks increasingly tend to use fewer routes. Across all the experiments in this section, the maximum $\tilde{q}$ observed was seven,

hence in practice, on-line route balancing problems are simpler than what might be imagined.

7.7 Discussion

The central idea of this chapter is to explore how using multiple routes in a wireless multi-hop network can increase the capacity. It was shown (Sec. 7.5) how to design wireless multi-hop networks that attain maximum possible capacity, as if there were no contention at all, by adding relay nodes and by dividing traffic over routes in a particularly structured way.

For random geometric networks (Sec. 7.6), the study solved for optimal capacity over the space of possible route assignments and found gains could often be achieved by using multiple routes for each origin-destination (end-to-end) pair. For larger networks, the gains are more dramatic and approach the theoretical contention-free limit. Closer examination of the optimal route assignments shows that they bear similarity to the designed networks of Sec. 7.5.

The capacity computations throughout the chapter are based on a theoretical capacity framework (Sec. 7.2) which is *exact*. (A tutorial example of these calculations is worked through in Sec. 7.3) It is well-known [11] that the intermediate steps in these kinds of calculations are NP-hard. However, it is shown in this chapter that the output is a relatively small system of constraints (size $\tilde{q}$) on route flows, and thus on-line capacity optimisation problems, that balance flow across routes as the demand vector varies, require only the solution of a relatively simple linear program. In principle, this approach seems tractable if the intensive component can be carried out off-line. Moreover, Gummadi et al. [103, 104] have proposed techniques for approximating the set of maximal co-transmitting sets that could in principle be incorporated in this framework to reduce the computational complexity of the off-line step.

Sec. 7.4 compared the theoretical capacity framework to simulation results. The comparison showed that potential constraints not captured by the framework may prevent a system achieving its maximum theoretical capacity. Specifically, a lack of queue management was shown to significantly reduce the achievable capacity region. However, when queues were managed, simulation results were comparable to analytically derived capacity regions, justifying the applicability of the Sec. 7.2 analytical method.

The results in Secs. 7.5 and 7.6 relate to wireless multi-hop networks with perfect beamforming and full-duplex capabilities — this is the best possible situation for capacity. Other experiments explored random geometric networks with beamforming capability alone, and found broadly similar trends, except, as might be expected, capacity is cut in half (see beamforming results in App. 7.A.1).

This Chapter assumes an idealised full-duplex system with perfect self-interference. When a node transmits and receives simultaneously, its own transmission imposes strong self-interference on its own reception. To remedy this problem, it is necessary to use a self-interference mitigation technique, which is fundamentally challenging to implement in a full-duplex radio. Various ap-

proaches have been investigated [155], including natural isolation (i.e., physical isolation between a nodes transmit and receive antennas) time-domain cancellation, (i.e., a node regenerates its self-interference signal, which is removed from the received signal in the time domain) and spatial suppression (i.e., using multiple transmit and receive antennas to suppress self-interference, in particular MIMO transmission schemes). The analysis of this chapter assumes an idealised full-duplex system with perfect self-interference capability, hence full-duplex in effect doubles the capacity of a single link. Of course, doubling the capacity of a link is not achievable and the gain from full-duplex is in fact less. The capacity regions derived throughout this chapter indicated a maximum theoretical upper bound. Applying a more realistic self-interference model, the volume of these regions would be reduced. Similarly, throughout the chapter, an idealised beamforming model is assumed. Applying a more realistic, imperfect, beamforming model would similarly reduce the achievable volume of the capacity region.

Sec. 7.3 showed a general progression in the capacity as advanced physical layer technologies are introduced. The framework could also be used to examine what gains can be achieved without beamforming or full-duplex, and how they depend on the relationship between TP and network density. In particular, the efficacy of adapting the CCA threshold or TP is moot [142].

7.8 Conclusions

This chapter contributes to knowledge in the following ways:

CTK 7.1 Demonstrating how to design networks to attain maximum possible capacity for multiple OD pairs.

CTK 7.2 Demonstrating that for complex multi-hop topologies, route constraints are much fewer in number than link constraints, and so dynamic route balancing is computationally feasible.

CTK 7.3 Providing statistics on the likelihood of capacity gains being achievable by intelligent routing in randomly generated multi-hop network topologies.

CTK 7.4 Identifying constraints preventing networks achieving the theoretical maximum capacity, specifically demonstrating the importance of queue management.

7.A Appendix

7.A.1 Beamforming Results

These tables relate to the analysis of ensembles of random networks as generated in Sec. 7.6. Compare with Tabs. 7.1 and 7.2, which relate to the application of beamforming and full-duplex together. In contrast these tables describe results when beamforming only is applied, and capacity gains are considerably less.

Table 7.3: Network statistics for experimental ensembles of random wireless multi-hop networks with beamforming capability only. The parameter m describes the number of relay nodes. Other parameters are as defined in Sec. 7.2. Results reported to three significant figures; integer results are exact.

m	$\langle n \rangle$	$\langle r \rangle$	$p_{\min}$	$\langle p \rangle$	$p_{\max}$	$\langle q \rangle$	$\langle \tilde{q} \rangle$	P($\tilde{q}$=1)
3	7	4	9	9	9	3	1	1
4	9.51	5.77	11	13.2	15	6.87	1	1
5	11.9	7.16	14	20.4	29	9.35	1.61	0.694
6	14.5	9.88	22	35.5	51	11.8	2.22	0.521
7	17.1	13.6	25	63.2	104	14.4	2.80	0.390
8	19.8	17.8	61	112	178	17.1	3.43	0.285

Table 7.4: Capacity statistics for experimental ensembles of random wireless multi-hop networks with beamforming capability only, characterised by maximising $d_{\mathrm{A}}+d_{\mathrm{B}}$ subject to $d_{\mathrm{A}} = d_{\mathrm{B}}$ (GAME 2). Here the superscript u refers to the number used in the optimal assignment and the subscript AB refers to the number shared in the optimal assignment. The parameter $\tilde{n} \leq n$ is the number of non-redundant links.

m	$\langle d_{\mathrm{A}} + d_{\mathrm{B}} \rangle$	$\langle r^{\mathrm{u}} \rangle$	P($r^{\mathrm{u}} = \tilde{q}+1$)	$\langle m^{\mathrm{u}}_{\mathrm{AB}} \rangle$	$\langle n^{\mathrm{u}}_{\mathrm{AB}}/\tilde{n} \rangle$
3	0.500	2	1	3	0.286
4	0.500	2	1	3.22	0.235
5	0.570	2.52	0.919	3.26	0.168
6	0.627	3.04	0.814	3.55	0.132
7	0.679	3.55	0.746	3.91	0.109
8	0.729	3.97	0.612	4.14	0.0849

CHAPTER 8

CONCLUSIONS

This thesis presented a framework for evaluating proposed technologies for next-generation wireless systems. The framework mixed hard systems modelling into a soft approach (see Fig. 1.1 p.5), providing a method for managing complexity and facilitating learning points for the development of future wireless systems. Here the approach and learning points are summarised (Sec 8.1), the context of the findings discussed (Sec. 8.2) and ideas for future work are proposed (Sec. 8.3).

8.1 Summary of Research and Original Contributions

The thesis began by exploration of the socio-economic system surrounding the physical communications infrastructure. A methodology for strategic decision making was proposed and used to identify key technologies of next-generation wireless networks likely to provide a return on research investments (Chapter 2). Further, existing methods of modelling the potential performance gains of proposed next-generation wireless technologies were reviewed and an opportunity for a new approach identified (Chapter 3). Subsequently, a novel conceptual simulation modelling methodology was proposed and developed (Chapter 4). The methodology was validated, showing that despite its significantly reduced complexity, it could be used to replicate results generated via normative simulation.

The methodology presented was used to quantitatively analyse and understand the performance impact of a selection of key next-generation wireless technologies. The thesis thus demonstrated the cyclical nature of a soft systems approach. With each application of the modelling methodology, tentative ideas for new applications emerged and informed the practice of how to apply the model to capture new situations of interest, characteristic of action research

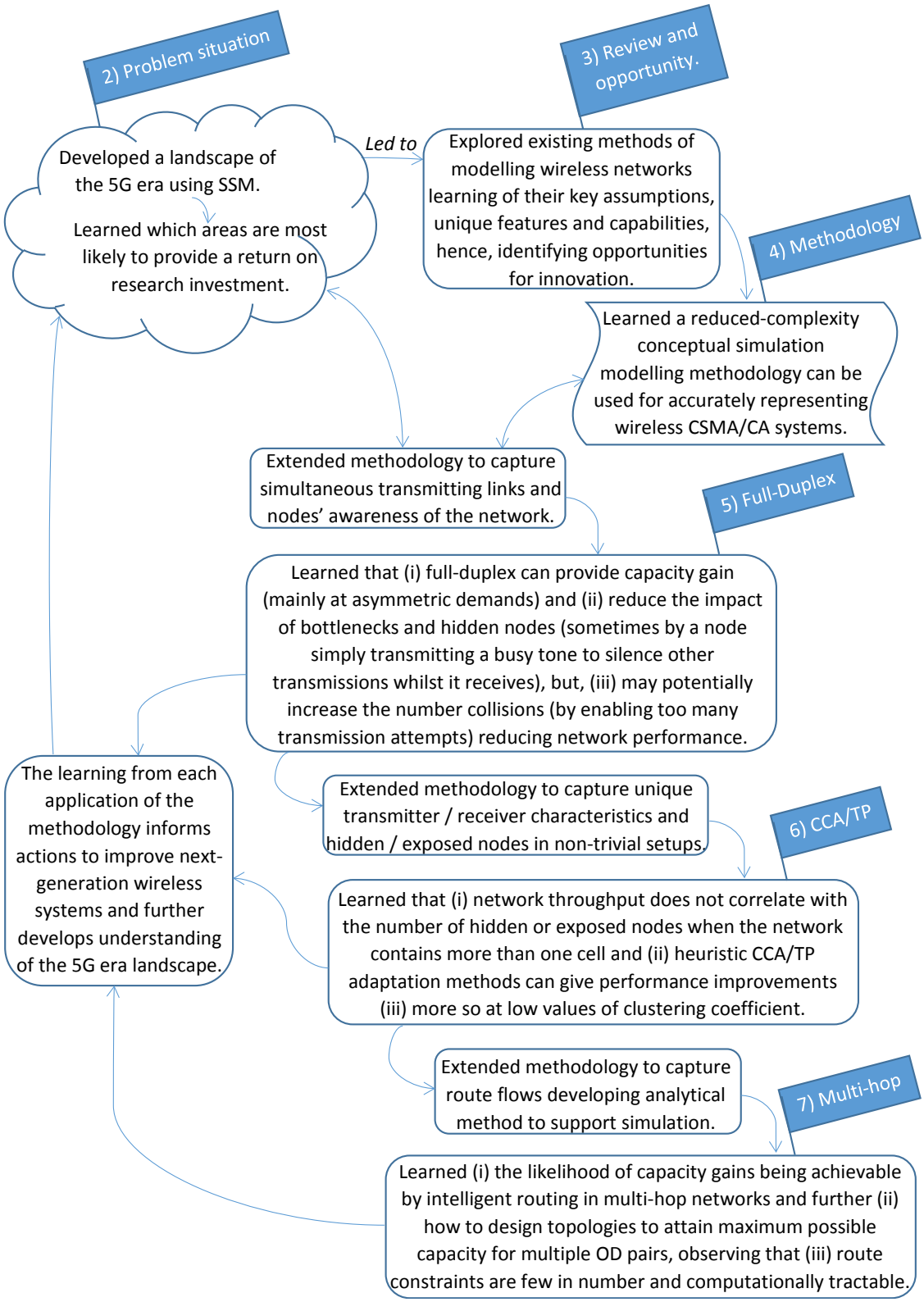


FIGURE 8.1. A development Fig. 1.1 (p.5) illustrating the cycle of inquiring and summarising the learning points generated in this thesis.

following a learning cycle [50].

Initially, the conceptual simulation modelling methodology was presented in its most basic format (Chapter 4). The simple situation of a wireless network where all links carrier sense and interfere with each other was discussed, capturing the interactions of those links with a 'collision' matrix. For its first application, the methodology was mapped to an investigation of full-duplex technology (Chapter 5). Here, the methodology was evolved to capture situations where multiple links in a network may co-transmit. Further, the model was extended to capture transmitters' and receivers' awareness of the network around them (or lack of awareness) with the introduction of the 'knowledge of the network' matrix. This matrix was used to capture hidden node situations in a simple network topology. From this, the capacity gain that full-duplex can provide was quantified, and shown to occur mainly at asymmetric demands. Further, full-duplex's ability to reduce the impact of bottlenecks and hidden nodes (sometimes by a node simply transmitting a busy tone to silence other transmissions whilst it receives) was demonstrated. A potential risk was also highlighted, with full-duplex in some cases enabling an increase in transmission attempts, which in turn led to a significant increase in the number collisions, resulting in an overall reduction in network performance.

The methodology was further evolved to capture the variation of nodes' TP and CCA threshold values (Chapter 6). With this, a method for exploring hidden and exposed node situations in non-trivial network set-ups was presented. It was shown that network throughput does not correlate with the number of hidden or exposed nodes when the network contains more than one cell. Further, simple heuristic CCA/TP adaptation methods can improve performance, especially when applied only to networks with low clustering coefficients.

Demonstrating another evolution of the methodology, and further developing an analytical method to support the simulation approach, multi-hop networks were explored (Chapter 7). Statistics on the likelihood of capacity gains being achievable by intelligent routing in multi-hop networks were provided, and further, it was demonstrated how topologies could be designed and managed to attain maximum possible capacity for multiple OD pairs.

Fig. 8.1 extends Fig. 1.1 (p.5) to reflect the cycle of inquiry in this thesis and to summarise the key learning points. A summary of the contributions to knowledge that emerged from each chapter is given in Appendix 8.A.1.

8.2 Overall Context of Main Findings

It is hoped that the modelling methodology demonstrated in this thesis can be used by other researchers in the wireless community to continue to study technologies for future CSMA/CA networks. More broadly, it is hoped that the principles of using hard systems modelling combined with a soft systems framework to facilitate learning, and further the principles of managing complexity by proposing a conceptual model, will be applied by other researchers modelling

different types of wireless systems or complex systems from other industries. It is hoped that the successful use of this multi-methodological approach will provide a useful reference to the systems engineering community.

The unique perspective, which the modelling approach provides, generated new knowledge as to the benefits, limitations, inefficiencies and 'trade-offs' of the proposed next-generation technologies. Despite the significantly reduced complexity approach, the modelling methodology was shown to be capable of producing results comparable to those produced via normative simulation. The simplification provides significant savings in set-up time (i.e., time to build the network of interest in the simulation environment) and allows the methodology to be easily mapped to scenarios of interest, enabling it to be used for exploring numerous design options as opposed to simply testing a given design.

In relation to the particular technologies investigated in this thesis, it was observed that, although when implemented alone in a network environment a capacity gain may be made, often that gain is much more significant when multiple technologies are implemented together or further an additional method of managing the technology is applied. For example, the benefits that full-duplex were able to provide were shown to be much more significant when combined with beamforming. Further, when adapting the transmission power of nodes in a network, the probability of this leading to a performance improvement is greater when only applied to networks with a low clustering coefficient. Additionally, implementing transmission power adaptation, beamforming or full-duplex in a multi-hop network fails to realise a significant capacity gain unless combined with appropriate queue management. The ease with which the simulation methodology can be mapped to different scenarios enabled these findings.

Statistics provided (throughout the thesis) in relation to the capacity region enable improved understanding of performance gains compared to the simple measurement of saturated flows. Visualising the capacity region facilitates an intuitive understanding of the efficiency and fairness of a system and further gives an impression of how the system performance differs at symmetric and asymmetric demand combinations.

In addition to answering a specific question each study set out to explore (such as does a particular technology add capacity to a network?), simulating an ensemble of demand combinations by scanning the capacity region enabled further findings. For example, it was observed that a full-duplex access point, with multiple clients associated to it, can mitigate against the problem of hidden nodes by simultaneously transmitting in full-duplex, or in the case that it has no data to transmit, simply transmitting noise to silence other clients in its cell. Further, it was observed that, although in some cases when adapting TP or CCA threshold a hidden node type situation may be introduced, this may actually increase the overall capacity of the system, as now, due to the interference pattern, if two transmissions collide one may always be successful while the other fails. Further still, exploring different route demand combinations in multi-hop networks identified that often the maximal throughput is achieved by splitting the demand unevenly across

a combination of routes. Throughout the investigations in this thesis, it was observed that often the most significant gains in capacity, realised from the introduction of new technologies, occurred at asymmetric demand combinations. None of these observations would have been made simply observing saturated or equal demand combinations as is common in much of the existing wireless literature.

8.3 Opportunities for Further Research

There are several opportunities for further research which build upon the contributions of this thesis. The simulation methodology presented has been used as a tool to learn about the potential performance benefits of just a select few next-generation technologies. There are, however, many other technologies proposed for the 5G era that the methodologies could be mapped to. The thesis only considered applications within a single channel, but the approach could easily be adapted to consider multiple channels or any situation where multiple queues are operating in parallel. For example, it seems the methodology would lend itself to modelling sending and receiving more than one data signal simultaneously via massive MIMO [92, 204]. Or further, the methodology could be used to capture how implementing a technology such as Mobile Edge Computing [158, 212], performing processing tasks closer to the edge of the network, could impact link data flows. Experiments could be designed to measure the impact of these on congestion in the network.

The approach of the thesis has been to consider networks designed to capture specific features of interest (e.g., hidden nodes) or to randomly generate ensembles of networks and extract statistics. Alternatively, however, the method could be used for simulation of 'bench test' networks more representative of the real world. Further, rather than mapping the methodology to investigate the performance impact of introducing new technologies into a network, it could instead be mapped to investigate the performance of, and subsequently optimise, networks in practical settings or potential applications of these networks. Examples may include, high-density networks in office blocks or shopping centres [72], or alternatively applications such as connected vehicles [159] or wireless backhaul systems [95].

The capacity region has been demonstrated as a valuable metric for generating an understanding of complex network interactions, however, performing ensemble simulations, as in this thesis, is limited by computational tractability. More sophisticated methods for searching high dimensional demand spaces, so that the capacity region can be identified, are necessary to extend the methodology to networks with higher numbers of links. In parallel, development of more sophisticated methods for visualising the capacity region in higher dimensions would be useful.

An alternative approach to exhaustive simulation is to consider (collections of) one-dimensional cuts through the demand space in the form $\mathbf{d} = d\mathbf{n}$ where d is total demand and $\mathbf{n}$ is a unit vector (in the one-norm). For understanding a simple capacity region (i.e., regular concave shape)

this can greatly reduce the number of simulations required. However, for a complex capacity region (i.e., an irregular shape) to be identified, a large number of vectors $\mathbf{n}$ must be considered, again leading to a large number of simulations. It is hence desirable to reduce the number of simulations required for each vector $\mathbf{n}$ to identify the boundary of the capacity region.

For a perfect M/M/1 queue, the point of saturation (i.e., demand needed to saturate the queue), and hence the capacity region boundary, could be identified for any single link by observation of its geometric queue length distribution [238] (assuming competing demands are fixed). Although similar, a queue on a node in a wireless network does not perform exactly in the manner of an M/M/1 queue, hence, alternative approaches to reducing the number of required simulations are needed. By observing the latency distribution, the queue length distribution and the inter-service times of a few transmissions, inside the capacity region, estimations can be made as to the location of the boundary.

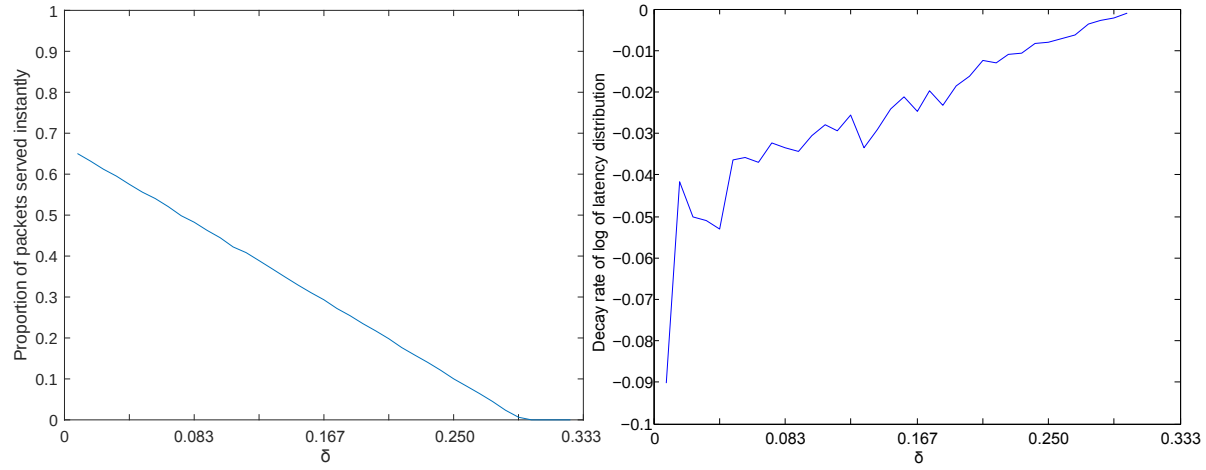
Consider the experimental set-up as follows. Three transmitting links are within carrier sensing range of each other: two links ($\mathscr{L}_1$ and $\mathscr{L}_2$) have fixed demand intensity (δ), the third link's ($\mathscr{L}_3$) δ is increased in 44 equal increments such that it saturates, resulting in 45 simulations. As throughout this thesis, the duration of each simulation is set at the time needed to transmit $10,000$ packets without collision.

It is possible to the observe, for $\mathscr{L}_3$, the proportion of packets arriving in the queue that are served instantly. Fig. 8.2(a) shows the proportion of packets served instantly appears to decrease, approximately linearly, with increasing δ. Assuming the relationship of the proportion of packets served instantly against δ is linear, extrapolating this line to cross the x-axis indicates the value of δ at which the capacity region boundary is reached.

Alternatively, consider the latency distribution of individual simulations within capacity. Taking log of the distribution at each δ and plotting the rate of decay of the geometric section against δ (see Fig. 8.2(b)), the plot can be extrapolated to estimate the capacity region boundary.

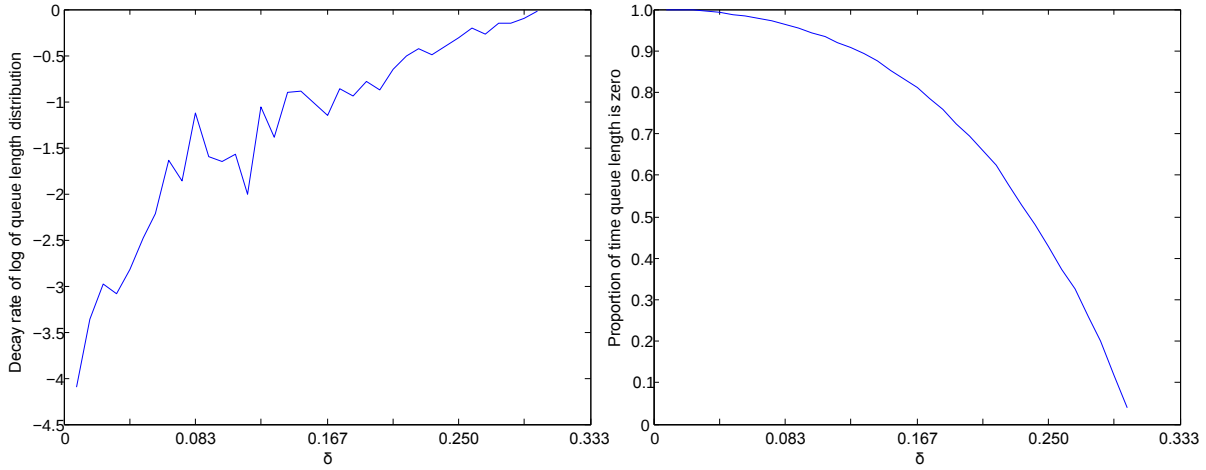
Observing the queue length distribution of $\mathscr{L}_3$ at a value of δ within the capacity region, a seemingly geometric distribution can be observed for a section of this simulation. Fig. 8.3(a) plots log of the geometric section of the queue length distribution's decay against simulated δ. Extrapolating this line to cross the x-axis indicates the value of δ that the capacity region is reached. Fig. 8.3(b) shows the proportion of time the queue length is zero for the same simulated values of δ. Again extrapolating this line to cross the x-axis indicates the value of δ at which the capacity region boundary is reached.

It is clear that for the methods described (shown in Figs. 8.2 and 8.3), the number of simulations across $\mathscr{L}_3$'s range of δ could be reduced (from 45), and still the capacity region boundary could be estimated. Of course, a higher number of simulations across the range, and a longer simulation time for each will increase the accuracy of the estimate. For each of these plots (particularly Figs. 8.2(b) and 8.3(a)), some oscillation in the plotted line is apparent and hence reducing the number significantly is likely to introduce inaccuracy.



(a) (b)

FIGURE 8.2. (a) Proportion of packets served instantly on arriving in the queue against δ. (b) Log of latency distributions geometric decay rate against δ. For both plots, extrapolating to the intersection with the x-axis can indicate the capacity region boundary and corresponding value of δ. For both fixed background traffic links, $\delta = 0.333$.



(a) (b)

FIGURE 8.3. (a) Log of queue length distribution decay against δ. (b) Proportion of time queue length is zero against δ. For both, extrapolating the plot to the intersection with the x-axis can indicate the capacity region boundary. For both fixed background traffic links, $\delta = 0.333$.

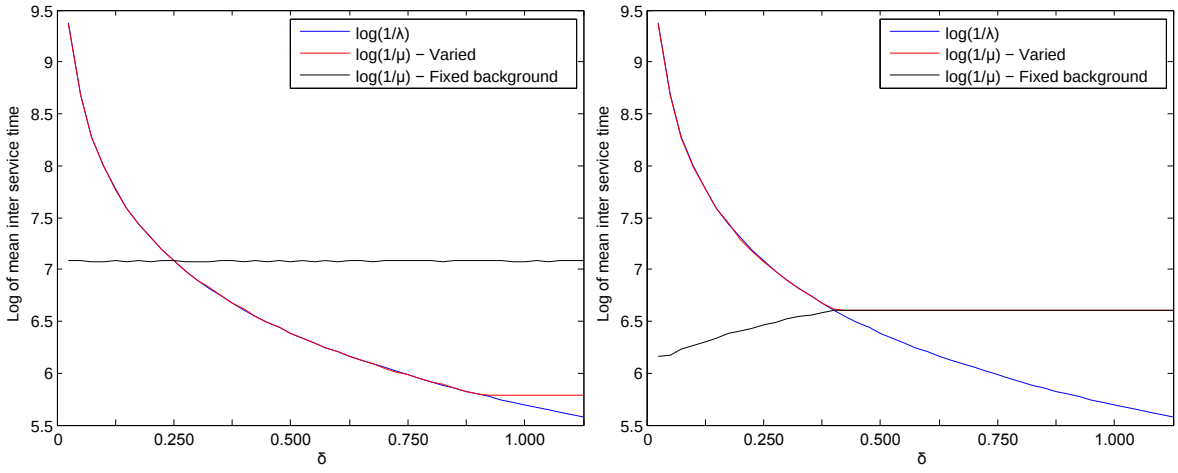


FIGURE 8.4. Plot of $\log(1/\mu)$ and $\log(1/\lambda)$ against δ. The point at which $\log(1/\mu)$ diverges from $\log(1/\lambda)$ indicates the δ at which $\mathscr{L}_3$ reaches the boundary of the capacity region. For both fixed background traffic links, (a) $\delta = 0.050$ and (b) $\delta = 0.600$

Another approach is to observe the inter-service time of $\mathscr{L}_3$'s transmissions. The mean inter-service times can be used to identify the boundary of the capacity region. Plotting $\log(1/\mu)$ (mean inter-service times is equal to $1/\mu$ where μ is the service rate) and $\log(1/\lambda)$ (where λ is the arrival rate) against δ indicates the demand intensity where the arrival rate is no longer being met as shown in Fig. 8.4. The figure presents two examples with different levels of background channel usage. In Fig. 8.4(a) the background channel demand is low (for each of the two transmitters $\delta = 0.050$) and in Fig. 8.4(b) the background is higher ($\delta = 0.600$). Because the CSMA/CA protocol divides the channel equally among transmitting links, within carrier sensing range of each other, when demand level is such that $\mathscr{L}_1$ and $\mathscr{L}_2$ are asking for less than their equal share of the medium, they will continue to satisfy their demand as δ to $\mathscr{L}_3$ increases (as Fig. 8.4(a)). When δ to $\mathscr{L}_3$ is sufficiently high such that it saturates, still queues at $\mathscr{L}_1$ and $\mathscr{L}_2$ will not build, requiring less than an equal share of the medium. When $\mathscr{L}_1$ and $\mathscr{L}_2$'s demand levels are higher than their equal share of the medium, they are forced to share the medium equally with $\mathscr{L}_3$ and hence their service rate decreases as δ to $\mathscr{L}_3$ increases (a queue builds on all transmitters).

The plotted $\log(1/\mu)$ is constant after the saturation point, i.e., the minimum inter-service time cannot be reduced further when adding additional demand beyond the saturation point. Because of this, the number of simulations across the range of δ could be reduced by simply simulating one link with a high δ, such that it is definitely saturated, and observing the mean inter-service time. This individual value $1/\mu$ can then be plotted as $\log(1/\mu)$ with $\log(1/\lambda)$. The point of saturation of the link can then be identified by extrapolating parallel to the x-axis from the plotted point $\log(1/\mu)$ to find the y-intercept of this horizontal line with $\log(1/\lambda)$. With

a sufficiently long simulation at fixed δ to all nodes, such that a high number of packets are transmitted, the mean inter-service time is very stable, and hence this method can significantly reduce the required number of simulations to identify the capacity region while maintaining a high level of accuracy.

However, the robust analysis and optimisation of these approaches requires specialist numerical analysis skills and are thus beyond the scope of this thesis and remain for future work.

Finally, a method for analytically computing the capacity region was presented in Chapter 7. The capacity computations throughout the chapter are based on a theoretical capacity framework which is *exact*. It is well-known [11] that the intermediate steps in these kinds of calculations are NP-hard, and hence for networks with a higher number of links, the computation is not tractable. Others [103, 104] have proposed techniques for approximating the set of maximal co-transmitting sets. Future work could explore incorporating these techniques into the framework of Chapter 7 to reduce the computational complexity. As was shown, post the intermediate steps, casting into route flows, the system of constraints reduces significantly. Hence, capacity optimisation problems balancing flow across routes as the demand vector varies, require only the solution of a relatively simple linear program. This approach seems tractable and given the reduction in the number of constraints, larger, potentially more interesting, network topologies could be investigated if the complexity of the intermediate steps could be reduced.

8.A Appendix

8.A.1 Summary of Contributions from Each Chapter

This thesis contributed to knowledge (CTK) in the following ways:

Chapter 2 by...

CTK 2.1 Presenting a methodology for developing a strategy for research investment in the complex development space of 5G era technologies.

CTK 2.2 Demonstrating how specific technical drivers of the 5G era are causing the evolution of the business infrastructure.

CTK 2.3 Demonstrating how problem structuring techniques can be used to develop a vision of the landscape of the 5G era for more informed decision making.

Chapter 3 by...

CTK 3.1 Defining conceptual simulation modelling, for the purpose of investigating wireless networks and identifying how it can be beneficial as a research methodology.

CTK 3.2 Identifying the key assumptions and implications of each of the three defined methodological categories: analytical modelling, normative simulation and conceptual simulation modelling, for the purpose of better-informing researchers in their selection.

CTK 3.3 Identifying the unique features and capabilities each methodology offers in the development of complex systems.

Chapter 4 by...

CTK 4.1 Presenting a valid conceptual simulation modelling methodology capable of representing wireless CSMA/CA networks.

CTK 4.2 Identifying that when considering zero carrier sensing error the capacity region is convex, and hence introducing new demands into a system increases the total system capacity. By increasing carrier sensing error (only slightly), the capacity region becomes concave, indicating a capacity penalty when a new demand is introduced. The concavity increases with increasing carrier sensing error.

Chapter 5 by...

CTK 5.1 Demonstrating the potential capacity gain that full-duplex can provide applied in various forms to a highly simplified mesh network set-up.

The first study in this chapter shows that full-duplex alone can increase the capacity of a network, however when full-duplex is combined with appropriate interference

management, the gain is much more significant. Existing work speculates at this result but this study is the first to quantify it.

CTK 5.2 Identifying that the majority of the capacity gain achieved by full-duplex occurs at asymmetric demand combinations.

This would not have been identified by just considering equal or saturated demands.

CTK 5.3 Showing that introducing full-duplex access points alone mitigates against the problem of bottlenecks, reduces the impact of hidden nodes and can increase the capacity of a network. It identifies that when full-duplex access points are able to work with full-duplex clients, the capacity gain is much more significant, however, it is shown that much of this capacity gain occurs at uneven demand combinations.

CTK 5.4 Identifying limitations of full-duplex. When the demand to all nodes is equally high, the introduction of full-duplex capability to clients is shown to increase the number of transmission attempts resulting in a significantly increased number of collisions and reduced network performance.

CTK 5.5 Discovering that at low traffic levels, a full-duplex access point may improve goodput by simply transmitting a busy tone to silence other transmissions whilst it receives, mitigating against the hidden node problem.

Chapter 6 by...

CTK 6.1 Presenting a modelling methodology for investigating CCA/TP adaptation.

CTK 6.2 Demonstrating heuristic CCA and TP adaptation methods that result in performance improvements.

CTK 6.3 Identifying that network throughput does not correlate with the number of hidden or exposed nodes when the network contains more than one cell and explaining how asymmetrically suppressing part of a network can remove competition for channel access and lead to overall network throughput improvement.

CTK 6.4 Identifying a relationship between throughput improvement and network clustering coefficient and thus network conditions under which CCA and TP adaptations are most likely to realise throughput improvements.

Chapter 7 by...

CTK 7.1 Demonstrating how to design networks to attain maximum possible capacity for multiple OD pairs.

CTK 7.2 Demonstrating that for complex multi-hop topologies, route constraints are much fewer in number than link constraints, and so dynamic route balancing is computationally feasible.

CTK 7.3 Providing statistics on the likelihood of capacity gains being achievable by intelligent routing in randomly generated multi-hop network topologies.

CTK 7.4 Identifying constraints preventing networks achieving the theoretical maximum capacity, specifically demonstrating the importance of queue management.

BIBLIOGRAPHY

[1] 5G-NOW, "5th Generation non-orthogonal waveforms for asynchronous waveforms," 2014. [Online]. Available: www.5gnow.eu

[2] 5G-XHaul, "5G-XHaul," 2014. [Online]. Available: http://www.5g-xhaul-project.eu/

[3] F. M. Abinader, E. P. Almeida, F. S. Chaves, A. M. Cavalcante, R. D. Vieira, R. C. Paiva, A. M. Sobrinho, S. Choudhury, E. Tuomaala, K. Doppler, and V. A. Sousa, "Enabling the coexistence of LTE and Wi-Fi in unlicensed bands," *IEEE Communications Magazine*, vol. 52, no. 11, pp. 54–61, 11 2014. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6957143

[4] F. Ackermann, "Problem structuring methods 'in the Dock': Arguing the case for Soft OR," *European Journal of Operational Research*, vol. 219, no. 3, pp. 652–658, 6 2012. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S0377221711010010

[5] M. S. Afaqui, E. Garcia-Villegas, E. Lopez-Aguilera, G. Smith, and D. Camps, "Evaluation of dynamic sensitivity control algorithm for IEEE 802.11ax," in *2015 IEEE Wireless Communications and Networking Conference (WCNC)*, no. Wcnc. IEEE, 3 2015, pp. 1060–1065. [Online]. Available: http://ieeexplore.ieee.org/document/7127616/

[6] I. Akyildiz and X. Wang, "A survey on wireless mesh networks," *IEEE Communications Magazine*, vol. 43, no. 9, pp. S23–S30, 9 2005. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1509968

[7] Alamsyah, M. H. Purnomo, I. K. E. Purnama, and E. Setijadi, "Performance of the routing protocols AODV, DSDV and OLSR in health monitoring using NS3," in *2016 International Seminar on Intelligent Technology and Its Applications (ISITIA)*. IEEE, 7 2016, pp. 323–328. [Online]. Available: http://ieeexplore.ieee.org/document/7828680/

[8] W. C. Anderton and M. Young, "Is Our Model for Contention Resolution Wrong?" 5 2017. [Online]. Available: http://arxiv.org/abs/1705.09271

[9] J. Andrews, S. Shakkottai, R. Heath, N. Jindal, M. Haenggi, R. Berry, D. Guo, M. Neely, S. Weber, S. Jafar, and A. Yener, "Rethinking information theory for mobile ad hoc networks," *IEEE Communications Magazine*, vol. 46, no. 12, pp. 94–101, 12 2008.

[Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4689214

[10] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, "What Will 5G Be?" *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1065–1082, 6 2014. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6824752

[11] E. Arikan, "Some Complexity Results about Packet Radio Networks," *Information Theory, IEEE*, vol. 30, no. 4, pp. 681–685, 1984. [Online]. Available: http://dx.doi.org/10.1109/TIT.1984.1056928

[12] B. Arthur, *The Nature of Technology: What it is and how it evolves*. Free Press, 2009.

[13] E. Aryafar, M. A. Khojastepour, K. Sundaresan, S. Rangarajan, and M. Chiang, "MIDU," in *Proceedings of the 18th annual international conference on Mobile computing and networking - Mobicom '12*. New York, New York, USA: ACM Press, 2012, p. 257. [Online]. Available: http://dl.acm.org/citation.cfm?doid=2348543.2348576

[14] P. Asthana, "Jumping the technology s-curve," *IEEE Spectrum*, vol. 32, no. 6, pp. 49–54, 6 1995. [Online]. Available: http://ieeexplore.ieee.org/document/387142/

[15] C. Avin, Z. Lotker, F. Pasquale, and Y.-A. Pignolet, "A Note on Uniform Power Connectivity in the SINR Model," in *Algorithmic Aspects Wireless Sensor Networks*, 6 2009, vol. 5804, pp. 116–127. [Online]. Available: http://link.springer.com/10.1007/978-3-642-05434-1_12

[16] F. Baccelli, P. Miihlethaler, and B. Blaszczyszyn, "Stochastic analysis of spatial and opportunistic aloha," *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 7, pp. 1105–1119, 9 2009. [Online]. Available: http://ieeexplore.ieee.org/document/5226963/

[17] F. Baccelli, "Stochastic Geometry and Wireless Networks: Volume II Applications," *Foundations and Trends® in Networking*, vol. 4, no. 1-2, pp. 1–312, 2009. [Online]. Available: http://www.nowpublishers.com/article/Details/NET-026

[18] A. Bahill and B. Gissing, "Re-evaluating systems engineering concepts using systems thinking," *IEEE Transactions on Systems, Man and Cybernetics, Part C (Applications and Reviews)*, vol. 28, no. 4, pp. 516–527, 1998. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=725338

[19] J. Bai, F. Qiu, and Y. Xue, "End-to-end fairness over non-convex capacity region in IEEE 802.11-based wireless networks," in *2012 IEEE 9th International Conference on Mobile Ad-Hoc and Sensor Systems (MASS 2012)*. IEEE, 10 2012, pp. 145–154. [Online]. Available: http://ieeexplore.ieee.org/document/6502512/

[20] S. Barghi, A. Khojastepour, K. Sundaresan, and S. Rangarajan, "Characterizing the throughput gain of single cell MIMO wireless systems with full duplex radios," *10th International Symposium on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt)*, no. May, pp. 68–74, 2012. [Online]. Available: http://ieeexplore.ieee.org/document/6260504/

[21] T. J. Barnett, A. Sumits, S. Jain, and U. Andra, "Cisco Visual Networking Index (VNI) Update Global Mobile Data Traffic Forecast," *Vni*, pp. 2015–2020, 2015. [Online]. Available: http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/complete-white-paper-c11-481360.html

[22] A. Barrat, M. Barthelemy, R. Pastor-Satorras, and A. Vespignani, "The architecture of complex weighted networks," *Proceedings of the National Academy of Sciences*, vol. 101, no. 11, pp. 3747–3752, 3 2004. [Online]. Available: http://www.pnas.org/cgi/doi/10.1073/pnas.0400087101

[23] B. Bellalta, "IEEE 802.11ax: High-efficiency WLANS," *IEEE Wireless Communications*, vol. 23, no. 1, pp. 38–46, 2 2016. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=7422404

[24] C. Bettstetter, "On the minimum node degree and connectivity of a wireless multihop network," in *Proceedings of the 3rd ACM international symposium on Mobile ad hoc networking & computing - MobiHoc '02*. New York, New York, USA: ACM Press, 2002, p. 80. [Online]. Available: http://portal.acm.org/citation.cfm?doid=513800.513811

[25] C. Bettstetter and C. Hartmann, "Connectivity of Wireless Multihop Networks in a Shadow Fading Environment," *Wireless Networks*, vol. 11, no. 5, pp. 571–579, 9 2005. [Online]. Available: http://link.springer.com/10.1007/s11276-005-3513-x

[26] D. Bharadia, E. McMilin, and S. Katti, "Full duplex radios," in *Proceedings of the ACM SIGCOMM 2013 conference on SIGCOMM - SIGCOMM '13*. New York, New York, USA: ACM Press, 2013, p. 375. [Online]. Available: http://dl.acm.org/citation.cfm?doid=2486001.2486033

[27] U. N. Bhat, *An Introduction to Queueing Theory*. Boston: Birkhäuser Boston, 2008. [Online]. Available: http://www.springerlink.com/index/10.1007/978-0-8176-4725-4

[28] G. Bianchi, A. Di Stefano, C. Giaconia, L. Scalia, G. Terrazzino, and I. Tinnirello, "Experimental Assessment of the Backoff Behavior of Commercial IEEE 802.11b Network Cards," in *IEEE INFOCOM 2007 - 26th IEEE International Conference on Computer Communications*. IEEE, 2007, pp. 1181–1189. [Online]. Available: http://ieeexplore.ieee.org/document/4215723/

[29] G. Bianchi, “Performance analysis of the IEEE 802.11 distributed coordination function,” *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 3, pp. 535–547, 3 2000. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=840210

[30] D. W. Bliss, T. M. Hancock, and P. Schniter, “Hardware phenomenological effects on cochannel full-duplex MIMO relay performance,” in *2012 Conference Record of the Forty Sixth Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*. IEEE, 11 2012, pp. 34–39. [Online]. Available: http://ieeexplore.ieee.org/document/6488953/

[31] D. W. Bliss, P. A. Parker, and A. R. Margetts, “Simultaneous Transmission and Reception for Improvedwireless Network Performance,” in *2007 IEEE/SP 14th Workshop on Statistical Signal Processing*. IEEE, 8 2007, pp. 478–482. [Online]. Available: http://ieeexplore.ieee.org/document/4301304/

[32] J. Boardman and B. Sauser, *Systemic Thinking, Building Maps For Worlds of Systems*. Wiley, 2013.

[33] R. Boorstyn, A. Kershenbaum, B. Maglaris, and V. Sahin, “Throughput Analysis in Multihop CSMA Packet Radio Networks,” *IEEE Transactions on Communications*, vol. 35, no. 3, pp. 267–274, 1987. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1096769

[34] K. E. Boulding, “General Systems Theory—The Skeleton of Science,” *Management Science*, vol. 2, no. 3, pp. 197–208, 4 1956. [Online]. Available: http://pubsonline.informs.org/doi/abs/10.1287/mnsc.2.3.197

[35] G. E. P. Box, “Science and Statistics,” *Journal of the American Statistical Association*, vol. 71, no. 356, pp. 791–799, 12 1976. [Online]. Available: http://www.tandfonline.com/doi/abs/10.1080/01621459.1976.10480949

[36] ——, “Robustness in the Strategy of Scientific Model Building,” in *Robustness in Statistics*. Elsevier, 1979, pp. 201–236. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/B9780124381506500182

[37] R. Bozidar, D. Gunawardena, A. Proutiere, N. Singh, V. Balan, and P. Key, “Efficiency and Fairness in Distributed Wireless Networks Through Self-interference Cancellation and Scheduling,” Microsoft Research, Cambridge, UK, Tech. Rep., 2009. [Online]. Available: https://www.microsoft.com/en-us/research/wp-content/uploads/2016/02/MSR-TR-2009-27.pdf

[38] D. Braess, “Über ein Paradoxon aus der Verkehrsplanung,” *Unternehmensforschung Operations Research - Recherche Opérationnelle*, vol. 12, no. 1, pp. 258–268, 12 1968. [Online]. Available: http://link.springer.com/10.1007/BF01918335

[39] J. M. Brazio, *Capacity analysis of multihop packet radio networks under a general class of channel access protocols and capture models*. Defense Technical Information Center, 1987. [Online]. Available: https://books.google.co.uk/books?id=Q_3zGwAACAAJ

[40] D. Bubley, "Disruptive Wireless: Thought-leading wireless industry analysis," 2014. [Online]. Available: http://disruptivewireless.blogspot.co.uk/2014/04/5g-standardisation-needs-to-be-multi.htm

[41] A. Busson and G. Chelius, "Point processes for interference modeling in CSMA/CA ad-hoc networks," in *Proceedings of the 6th ACM symposium on Performance evaluation of wireless ad hoc, sensor, and ubiquitous networks - PE-WASUN '09*. New York, New York, USA: ACM Press, 2009, p. 33. [Online]. Available: http://portal.acm.org/citation.cfm?doid=1641876.1641884

[42] A. Busson, G. Chelius, and J.-M. Gorce, "Interference Modeling in CSMA Multi-Hop Wireless Networks," *INRIA*, p. 21, 2009. [Online]. Available: http://hal.inria.fr/inria-00316029

[43] F. Cao, Z. Zhong, Z. Fan, M. Sooriyabandara, S. Armour, and A. Ganesh, "User association for load balancing with uneven user distribution in IEEE 802.11ax networks," in *2016 13th IEEE Annual Consumer Communications & Networking Conference (CCNC)*. IEEE, 1 2016, pp. 487–490. [Online]. Available: http://ieeexplore.ieee.org/document/7444828/

[44] J. Carson, "Convincing Users of Model's Validity is Challenging Aspect of Modeler's Job," *Industrial Engineering*, vol. 18, no. 6, pp. 74–85, 1986.

[45] CCHP, "Center For Connected Healthcare Policy - The National Telehealth Policy Resource Centre." [Online]. Available: http://www.cchpca.org/what-is-telehealth

[46] CfPH, "Centre for Pervasive Healthcare." [Online]. Available: http://www.pervasivehealthcare.dk/

[47] H. Chang, V. Misra, and D. Rubenstein, "A General Model and Analysis of Physical Layer Capture in 802.11 Networks," in *Proceedings IEEE INFOCOM 2006. 25TH IEEE International Conference on Computer Communications*, vol. 00, no. c. IEEE, 2006, pp. 1–12. [Online]. Available: http://ieeexplore.ieee.org/document/4146871/

[48] P. Checkland, *Systems thinking, systems practice*. Cumincad, 1981.

[49] ——, *Systems Thinking, Systems Practice (30 Year Retrospective)*. Oxford UK: Wiley, 1999.

[50] ——, "Soft Systems Methodology : A Thirty Year Retrospective," *Systems Research and Behavioral Science*, vol. 17, no. 1, pp. 11–58, 2000. [Online].

Available: http://onlinelibrary.wiley.com/doi/10.1002/1099-1743(200011)17:1+%3C::AID-SRES374%3E3.0.CO;2-O/abstract

[51] S. Chen, M. Beach, and J. McGeehan, "Division-free duplex for wireless applications," *Electronics Letters*, vol. 34, no. 2, p. 147, 1998. [Online]. Available: http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=653163

[52] W. Cheng, X. Zhang, and H. Zhang, "RTS/FCTS mechanism based full-duplex MAC protocol for wireless networks," in *2013 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 12 2013, pp. 5017–5022. [Online]. Available: http://ieeexplore.ieee.org/document/6855746/

[53] Y.-C. Cheng and T. Robertazzi, "Critical connectivity phenomena in multihop radio models," *IEEE Transactions on Communications*, vol. 37, no. 7, pp. 770–777, 7 1989. [Online]. Available: http://ieeexplore.ieee.org/document/31170/

[54] G. Chengetanai and G. B. O'Reilly, "Survey on simulation tools for wireless mobile ad hoc networks," in *2015 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT)*. IEEE, 3 2015, pp. 1–7. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=7226167

[55] H. Chesbrough, "Business model innovation: it's not just about technology anymore," *Strategy & Leadership*, vol. 35, no. 6, pp. 12–17, 2007. [Online]. Available: http://www.emeraldinsight.com/10.1108/10878570710833714

[56] M. Chiang, P. Hande, T. Lan, and C. W. Tan, "Power Control in Wireless Cellular Networks," *Foundations and Trends® in Networking*, vol. 2, no. 4, pp. 381–533, 2007. [Online]. Available: http://www.nowpublishers.com/article/Details/NET-009

[57] China-Mobile, "5G White paper China Mobile," vol. 96, pp. 3S–4S, 2015. [Online]. Available: http://www.wendangku.net/doc/3148b49efad6195f302ba644.html

[58] J. I. Choi, M. Jain, K. Srinivasan, P. Levis, and S. Katti, "Achieving single channel, full duplex wireless communication," in *Proceedings of the sixteenth annual international conference on Mobile computing and networking - MobiCom '10*. New York, New York, USA: ACM Press, 2010, p. 1. [Online]. Available: http://portal.acm.org/citation.cfm?doid=1859995.1859997

[59] T. S. Chouhan and R. S. Deshmukh, "Analysis of DSDV, OLSR and AODV Routing Protocols in VANETS Scenario: Using NS3," in *2015 International Conference on Computational Intelligence and Communication Networks (CICN)*. IEEE, 12 2015, pp. 85–89. [Online]. Available: http://ieeexplore.ieee.org/document/7546061/

[60] Cisco, "Cisco Visual Networking Index : Global Mobile Data Traffic Forecast Update , 2013 - 2018," *White Paper*, pp. 2013–2018, 2014. [Online]. Available: http://www.anatel.org.mx/docs/interes/Cisco_VNI_Forecast_and_Methodology.pdf

[61] W. M. Cohen and D. A. Levinthal, "Innovation and Learning: The Two Faces of R&D," *The Economic Journal*, vol. 99, no. 397, p. 569, 9 1989. [Online]. Available: https://www.jstor.org/stable/2233763?origin=crossref

[62] D. R. Cox, "Some Statistical Methods Connected with Series of Events," *Journal of the Royal Statistical Society*, vol. 17, no. 2, pp. 129–164, 1955. [Online]. Available: https://www.nuffield.ox.ac.uk/users/cox/cox36.pdf

[63] D. R. Cox and W. Smith, *Queues*, ser. Chapman & Hall/CRC Monographs on Statistics & Applied Probability. Taylor & Francis, 1991. [Online]. Available: https://books.google.co.uk/books?id=ZP9Rg_dx_JUC

[64] R. Cruz and A. Santhanam, "Optimal routing, link scheduling and power control in multihop wireless networks," in *IEEE INFOCOM 2003. Twenty-second Annual Joint Conference of the IEEE Computer and Communications Societies (IEEE Cat. No.03CH37428)*, vol. 1, no. C. IEEE, 2003, pp. 702–711. [Online]. Available: http://ieeexplore.ieee.org/document/1208720/

[65] S. M. Das, D. Koutsonikolas, Y. C. Hu, and D. Peroulis, "Characterizing multi-way interference in wireless mesh networks," in *Proceedings of the 1st international workshop on Wireless network testbeds, experimental evaluation & characterization - WiNTECH '06*. New York, New York, USA: ACM Press, 2006, p. 57. [Online]. Available: http://portal.acm.org/citation.cfm?doid=1160987.1160999

[66] J. Davis, A. MacDonald, and L. White, "Problem-structuring methods and project management: an example of stakeholder involvement using Hierarchical Process Modelling methodology," *Journal of the Operational Research Society*, vol. 61, no. S6, pp. 893–904, 6 2010. [Online]. Available: http://link.springer.com/10.1057/jors.2010.12

[67] J. P. Davis and J. W. Hall, "A software-supported process for assembling evidence and handling uncertainty in decision-making," *Decision Support Systems*, vol. 35, no. 3, pp. 415–433, 2003. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S0167923602001173

[68] J. Davis, A. MacDonald, and S. Marashi, "Integrated performance measurement to support strategic decision making in engineering organisations," in *2007 IEEE International Conference on Industrial Engineering and Engineering Management*. IEEE, 12 2007, pp. 1762–1766. [Online]. Available: http://ieeexplore.ieee.org/document/4419495/

[69] P. Davis, "Generalizing concepts of verification, validation and accreditation (VV&A) for military simulation," Santa Monica, CA, 1992. [Online]. Available: https://www.rand.org/pubs/reports/R4249.html

[70] B. P. Day, A. R. Margetts, D. W. Bliss, and P. Schniter, "Full-Duplex Bidirectional MIMO: Achievable Rates Under Limited Dynamic Range," *IEEE Transactions on Signal Processing*, vol. 60, no. 7, pp. 3702–3713, 7 2012. [Online]. Available: http://ieeexplore.ieee.org/document/6177689/

[71] Defense_Acquisition_Univ_FT_Belvoir_VA, *System Engineering Fundamentals*. VIRGINIA: Defense Acquisition University Press Fort Belvoir, 2001. [Online]. Available: http://www.dtic.mil/dtic/tr/fulltext/u2/a606327.pdf

[72] D.-J. Deng, K.-C. Chen, and R.-S. Cheng, "IEEE 802.11ax: Next generation wireless local area networks," in *10th International Conference on Heterogeneous Networking for Quality, Reliability, Security and Robustness*, vol. 1. IEEE, 8 2014, pp. 77–82. [Online]. Available: http://ieeexplore.ieee.org/document/6928663/

[73] J. Deng, B. Liang, and P. Varshney, "Tuning the carrier sensing range of IEEE 802.11 MAC," in *IEEE Global Telecommunications Conference, 2004. GLOBECOM '04.*, vol. 5. IEEE, 2004, pp. 2987–2991. [Online]. Available: http://ieeexplore.ieee.org/document/1378900/

[74] M. Desai and D. Manjunath, "On the connectivity in finite ad hoc networks," *IEEE Communications Letters*, vol. 6, no. 10, pp. 437–439, 10 2002. [Online]. Available: http://ieeexplore.ieee.org/document/1042240/

[75] A. Doufexi, "Vehicular Communications in TV Whitespaces," 2015. [Online]. Available: http://www.bris.ac.uk/engineering/research/csn/projects/past/vehicularcommunications.html

[76] O. Dousse, F. Baccelli, and P. Thiran, "Impact of interferences on connectivity in ad hoc networks," *IEEE/ACM Transactions on Networking*, vol. 13, no. 2, pp. 425–436, 4 2005. [Online]. Available: http://ieeexplore.ieee.org/document/1424049/

[77] O. Dousse, P. Thiran, and M. Durvy, "On the fairness of large CSMA networks," *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 7, pp. 1093–1104, 9 2009. [Online]. Available: http://ieeexplore.ieee.org/document/5226962/

[78] M. Duarte, C. Dick, and A. Sabharwal, "Experiment-Driven Characterization of Full-Duplex Wireless Systems," *IEEE Transactions on Wireless Communications*, vol. 11, no. 12, pp. 4296–4307, 12 2012. [Online]. Available: http://ieeexplore.ieee.org/document/6489234/

[79] M. Duarte and A. Sabharwal, "Full-duplex wireless communications using off-the-shelf radios: Feasibility and first results," in *2010 Conference Record of the Forty Fourth Asilomar Conference on Signals, Systems and Computers*. IEEE, 11 2010, pp. 1558–1562. [Online]. Available: http://ieeexplore.ieee.org/document/5757799/

[80] J. Epstein, "Why Model?" *Cybernetics and Systems*, vol. 35, no. 2-3, pp. 117–128, 3 2008. [Online]. Available: http://www.tandfonline.com/doi/abs/10.1080/01969720490426803

[81] B. G. Evans, "The role of satellites in 5G," *2014 7th Advanced Satellite Multimedia Systems Conference and the 13th Signal Processing for Space Communications Workshop (ASMS/SPSC)*, pp. 197–202, 9 2014. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6934544

[82] E. Everett, M. Duarte, C. Dick, and A. Sabharwal, "Empowering full-duplex wireless communication by exploiting directional diversity," in *2011 Conference Record of the Forty Fifth Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*. IEEE, 11 2011, pp. 2002–2006. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6190376

[83] FED4FIRE, "FED4FIRE." [Online]. Available: https://www.fed4fire.eu/

[84] C. Felita and M. Suryanegara, "5G key technologies: Identifying innovation opportunity," *2013 International Conference on QiR*, pp. 235–238, 6 2013. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6632571

[85] FIBRE, "FIBRE." [Online]. Available: http://www.fibre-ict.eu/

[86] G. S. Fishman, *Discrete-Event Simulation; Modeling, Programming and Analysis*. Springer, 2001.

[87] Flourish, "Flourish." [Online]. Available: http://www.flourishmobility.com/

[88] H. Friis, "A Note on a Simple Transmission Formula," *Proceedings of the IRE*, vol. 34, no. 5, pp. 254–256, 5 1946. [Online]. Available: http://ieeexplore.ieee.org/document/1697062/

[89] A. Frotzscher and G. Fettweis, "Digital compensation of transmitter leakage in FDD zero-IF receivers," *Transactions on Emerging Telecommunications Technologies*, vol. 23, no. 2, pp. 105–120, 3 2012. [Online]. Available: http://doi.wiley.com/10.1002/ett.1514

[90] FUTEBOL, "FUTEBOL - Federated Union of Telecommunications Research Facilities for an EU-Brazil Open Laboratory." [Online]. Available: http://www.ict-futebol.org.br/

[91] Y. Gao, D.-M. Chiu, and J. C. Lui, "Determining the end-to-end throughput capacity in multi-hop networks," *ACM SIGMETRICS Performance Evaluation Review*, vol. 34, no. 1, p. 39, 6 2006. [Online]. Available: http://portal.acm.org/citation.cfm?doid=1140103.1140284

[92] Z. Gao, L. Dai, D. Mi, Z. Wang, M. A. Imran, and M. Z. Shakir, "MmWave massive-MIMO-based wireless backhaul for the 5G ultra-dense network," *IEEE Wireless Communications*, vol. 22, no. 5, pp. 13–21, 10 2015. [Online]. Available: http://ieeexplore.ieee.org/document/7306533/

[93] M. Garetto, T. Salonidis, and E. Knightly, "Modeling Per-Flow Throughput and Capturing Starvation in CSMA Multi-Hop Wireless Networks," *IEEE/ACM Transactions on Networking*, vol. 16, no. 4, pp. 864–877, 8 2008. [Online]. Available: http://ieeexplore.ieee.org/document/4457986/

[94] Gateway-to-Research, "Pathway to Autonomous Commercial Vehicles." [Online]. Available: http://gtr.rcuk.ac.uk/projects?ref=102585

[95] X. Ge, H. Cheng, M. Guizani, and T. Han, "5G wireless backhaul networks: challenges and research advances," *IEEE Network*, vol. 28, no. 6, pp. 6–11, 11 2014. [Online]. Available: http://ieeexplore.ieee.org/document/6963798/

[96] GEANT, "GEANT." [Online]. Available: https://www.geant.org/

[97] A. Gohil, H. Modi, and S. K. Patel, "5G technology of mobile communication: A survey," *2013 International Conference on Intelligent Systems and Signal Processing (ISSP)*, pp. 288–292, 3 2013. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6526920

[98] S. Gold, "5G is the communications answer, now what is the question?" 2014. [Online]. Available: url:http://www.theguardian.com/innovation-nation-awards/5g-is-the-communications-answer-now-what-is-the-question

[99] Google, "Facts about Googles acquisition of Motorola," 2014. [Online]. Available: https://www.google.com/press/motorola/

[100] ——, "Google Self-Driving Car Project," 2015. [Online]. Available: https://www.google.com/selfdrivingcar/

[101] A. A. Goulianos, N. F. Abdullah, D. Kong, E. Mellios, D. Berkovskyy, A. Doufexi, and A. Nix, "Evaluation of 802.11 and LTE for Automotive Applications," in *2014 IEEE 80th Vehicular Technology Conference (VTC2014-Fall)*. IEEE, 9 2014, pp. 1–5. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6965970

[102] S. Goyal, P. Liu, S. S. Panwar, R. A. Difazio, R. Yang, and E. Bala, "Full duplex cellular systems: will doubling interference prevent doubling capacity?" *IEEE Communications Magazine*, vol. 53, no. 5, pp. 121–127, 5 2015. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=7105650

[103] R. Gummadi, K. Jung, D. Shah, and R. Sreenivas, "Feasible Rate Allocation in Wireless Networks," in *IEEE INFOCOM 2008 - The 27th Conference on Computer Communications*. IEEE, 4 2008, pp. 995–1003. [Online]. Available: http://ieeexplore.ieee.org/document/4509748/

[104] ——, "Computing the Capacity Region of a Wireless Network," in *IEEE INFOCOM 2009 - The 28th Conference on Computer Communications*. IEEE, 4 2009, pp. 1341–1349. [Online]. Available: http://ieeexplore.ieee.org/document/5062049/

[105] L. Guo, K. Long, K. Munasinghe, and X. Wei, "Multi-hop relaying in 5G: From research to systems, standards, and applications," *China Communications*, vol. 13, no. 10, pp. iii–iv, 10 2016. [Online]. Available: http://ieeexplore.ieee.org/document/7732006/

[106] P. Gupta and P. R. Kumar, "Critical Power for Asymptotic Connectivity in Wireless Networks," in *Stochastic Analysis, Control, Optimization and Applications SE - 33*, ser. Systems & Control: Foundations & Applications, W. McEneaney, G. Yin, and Q. Zhang, Eds. Birkhäuser Boston, 1999, pp. 547–566. [Online]. Available: http://dx.doi.org/10.1007/978-1-4612-1784-8_33

[107] ——, "The capacity of wireless networks," *IEEE Transactions on Information Theory*, vol. 46, no. 2, pp. 388–404, 3 2000. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=825799

[108] M. Haenggi, J. Andrews, F. Baccelli, O. Dousse, and M. Franceschetti, "Stochastic geometry and random graphs for the analysis and design of wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 7, pp. 1029–1046, 9 2009. [Online]. Available: http://ieeexplore.ieee.org/document/5226957/

[109] M. Haenggi, "Mean Interference in Hard-Core Wireless Networks," *IEEE Communications Letters*, vol. 15, no. 8, pp. 792–794, 8 2011. [Online]. Available: http://ieeexplore.ieee.org/document/5934671/

[110] ——, *Stochastic Geometry for Wireless Networks*. Cambridge: Cambridge University Press, 2012. [Online]. Available: http://ebooks.cambridge.org/ref/id/CBO9781139043816

[111] J. W. Hall, D. I. Blockley, and J. P. Davis, "Uncertain inference using interval probability theory," *International Journal of Approximate Reasoning*, vol. 19, no. 3-4, pp. 247–264, 10 1998. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S0888613X98100105

[112] A. Hasan and J. Andrews, "The Guard Zone in Wireless Ad hoc Networks," *IEEE Transactions on Wireless Communications*, vol. 6, no. 3, pp. 897–906, 3 2007. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4133877

[113] R. Hekmat and P. Van Mieghem, "Connectivity in Wireless Ad-hoc Networks with a Log-normal Radio Model," *Mobile Networks and Applications*, vol. 11, no. 3, pp. 351–360, 6 2006. [Online]. Available: http://link.springer.com/10.1007/s11036-006-5188-7

[114] D. Helbing, "From microscopic to macroscopic traffic models," in *A Perspective Look at Nonlinear Media*. Springer Berlin Heidelberg, 1998, pp. 122–139. [Online]. Available: http://www.springerlink.com/index/10.1007/BFb0104959

[115] Horizon-2020, "Horizon 2020 Work Program 2016-2017," 2016. [Online]. Available: http://ec.europa.eu/research/participants/data/ref/h2020/other/wp/2016-2017/annexes/h2020-wp1617-annex-ga_en.pdf

[116] E. Hossain, M. Rasti, H. Tabassum, and A. Abdelnasser, "Evolution toward 5G multi-tier cellular wireless networks: An interference management perspective," *IEEE Wireless Communications*, vol. 21, no. 3, pp. 118–127, 6 2014. [Online]. Available: http://ieeexplore.ieee.org/document/6845056/

[117] K.-C. Huang and K.-C. Chen, "Interference analysis of nonpersistent CSMA with hidden terminals in multicell wireless data networks," in *Proceedings of 6th International Symposium on Personal, Indoor and Mobile Radio Communications*, vol. 2. IEEE, 1995, pp. 907–911. [Online]. Available: http://ieeexplore.ieee.org/document/480998/

[118] I-motors, "i-motors." [Online]. Available: http://www.i-motors.cloud/

[119] IEEE, "IEEE Standard for Information Technology - Telecommunications and Information Exchange Between Systems - Local and Metropolitan Area Networks - Specific Requirements - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specificatio," pp. 1–1076, 2007.

[120] INTACT, "INTACT." [Online]. Available: http://gtr.rcuk.ac.uk/projects?ref=102587

[121] R. Ison, "Systems Thinking and Practice for Action Research," in *The SAGE Handbook of Action Research*. 1 Oliver's Yard, 55 City Road, London England EC1Y 1SP United Kingdom: SAGE Publications Ltd, 2008, pp. 139–158. [Online]. Available: http://methods.sagepub.com/book/the-sage-handbook-of-action-research/d15.xml

[122] A. Jain, A. K. Dubey, R. Upadhyay, and S. Charhate, "Performance Evaluation of Wireless Network in Presence of Hidden Node: A Queuing Theory Approach," in *2008 Second Asia International Conference on Modelling & Simulation (AMS)*. IEEE, 5 2008, pp. 225–229. [Online]. Available: http://ieeexplore.ieee.org/document/4530480/

[123] K. Jain, J. Padhye, V. N. Padmanabhan, and L. Qiu, "Impact of interference on multi-hop wireless network performance," *Proceedings of the 9th annual international conference*

on Mobile computing and networking - MobiCom '03, p. 66, 2003. [Online]. Available: http://portal.acm.org/citation.cfm?doid=938985.938993

[124] M. Jain, J. I. Choi, T. Kim, D. Bharadia, S. Seth, K. Srinivasan, P. Levis, S. Katti, and P. Sinha, "Practical, real-time, full duplex wireless," in *Proceedings of the 17th annual international conference on Mobile computing and networking - MobiCom '11*. New York, New York, USA: ACM Press, 2011, p. 301. [Online]. Available: http://dl.acm.org/citation.cfm?doid=2030613.2030647

[125] R. Jain, A. Durresi, and G. Babic, "Throughput fairness index: An explanation," Tech. Rep., 1999. [Online]. Available: http://www.cse.wustl.edu/~jain/atmf/ftp/atm99-0045.pdf

[126] J. Jeong, H. Kim, S. Lee, and J. Shin, "An analysis of hidden node problem in IEEE 802.11 multihop networks," in *Sixth International Conference on Networked Computing and Advanced Information Management (NCM)*, 2010, pp. 282–285. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5573151

[127] L. B. Jiang and S. C. Liew, "Improving Throughput and Fairness by Reducing Exposed and Hidden Nodes in 802.11 Networks," *IEEE Transactions on Mobile Computing*, vol. 7, no. 1, pp. 34–49, 1 2008. [Online]. Available: http://ieeexplore.ieee.org/document/4359012/

[128] A. Jindal and K. Psounis, "Achievable rate region and optimality of multi-hop wireless 802.11-scheduled networks," in *2008 Information Theory and Applications Workshop*. IEEE, 1 2008, pp. 263–269. [Online]. Available: http://ieeexplore.ieee.org/document/4601059/

[129] ——, "The Achievable Rate Region of 802.11-Scheduled Multihop Networks," *IEEE/ACM Transactions on Networking*, vol. 17, no. 4, pp. 1118–1131, 8 2009. [Online]. Available: http://ieeexplore.ieee.org/document/4982658/

[130] ——, "On the Efficiency of CSMA-CA Scheduling in Wireless Multihop Networks," *IEEE/ACM Transactions on Networking*, vol. 21, no. 5, pp. 1392–1406, 10 2013. [Online]. Available: http://ieeexplore.ieee.org/document/6361253/

[131] R. B. Johnson and A. J. Onwuegbuzie, "Mixed Methods Research: A Research Paradigm Whose Time Has Come," *Educational Researcher*, vol. 33, no. 7, pp. 14–26, 10 2004. [Online]. Available: http://journals.sagepub.com/doi/10.3102/0013189X033007014

[132] W. Jones, M. Sooriyabandara, M. Yearworth, A. Doufexi, and R. E. Wilson, "Planning For 5G: A Problem Structuring Approach for Survival in the Telecoms Industry," *Systems Engineering*, vol. 19, no. 4, pp. 301–321, 7 2016. [Online]. Available: http://doi.wiley.com/10.1002/sys.21354

[133] W. Jones, R. E. Wilson, M. Sooriyabandara, and A. Doufexi, "Wireless Network MAC Layer Performance Evaluation with Full-Duplex Capable Nodes," in *Proceedings of the 12th ACM Symposium on QoS and Security for Wireless and Mobile Networks - Q2SWinet '16*. New York, New York, USA: ACM Press, 2016, pp. 111–118. [Online]. Available: http://dl.acm.org/citation.cfm?doid=2988272.2990294

[134] H. Ju, S. Lim, D. Kim, H. V. Poor, and D. Hong, "Full Duplexity in Beamforming-Based Multi-Hop Relay Networks," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 8, pp. 1554–1565, 9 2012. [Online]. Available: http://ieeexplore.ieee.org/document/6280259/

[135] C. Kai and S. Zhang, "Throughput analysis of CSMA wireless networks with finite offered-load," in *2013 IEEE International Conference on Communications (ICC)*. IEEE, 6 2013, pp. 6101–6106. [Online]. Available: http://ieeexplore.ieee.org/document/6655579/

[136] S. Kannangara and M. Faulkner, "Analysis of an Adaptive Wideband Duplexer With Double-Loop Cancellation," *IEEE Transactions on Vehicular Technology*, vol. 56, no. 4, pp. 1971–1982, 7 2007. [Online]. Available: http://ieeexplore.ieee.org/document/4277076/

[137] W. Kasch, J. Ward, and J. Andrusenko, "Wireless network modeling and simulation tools for designers and developers," *IEEE Communications Magazine*, vol. 47, no. 3, pp. 120–127, 3 2009. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4804397

[138] F. E. Kast and J. E. Rosenzweig, "General System Theory: Applications for Organization and Management." *Academy of Management Journal*, vol. 15, no. 4, pp. 447–465, 12 1972. [Online]. Available: http://amj.aom.org/cgi/doi/10.2307/255141

[139] F. P. Kelly, "The Clifford Paterson Lecture, 1995 Modelling communication networks, present and future," *Philosophical Transactions: Mathematical, Physical and Engineering Sciences*, vol. 354, no. 1707, pp. 437–463, 1996. [Online]. Available: http://www.jstor.org/stable/54588

[140] M. A. Khojastepour and S. Rangarajan, "Wideband digital cancellation for full-duplex communications," in *2012 Conference Record of the Forty Sixth Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*. IEEE, 11 2012, pp. 1300–1304. [Online]. Available: http://ieeexplore.ieee.org/document/6489234/

[141] E. Khorov, A. Kiryanov, and A. Lyakhov, "IEEE 802.11ax: How to Build High Efficiency WLANs," in *2015 International Conference on Engineering and Telecommunication (EnT)*. IEEE, 11 2015, pp. 14–19. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=7420884

[142] T.-S. Kim, J. C. Hou, and H. Lim, "Improving spatial reuse through tuning transmit power, carrier sense threshold, and data rate in multihop wireless networks," *Proceedings of the 12th annual international conference on Mobile computing and networking - MobiCom '06*, p. 366, 2006. [Online]. Available: http://portal.acm.org/citation.cfm?doid=1161089.1161131

[143] L. Kleinrock and F. Tobagi, "Packet Switching in Radio Channels: Part I–Carrier Sense Multiple-Access Modes and Their Throughput-Delay Characteristics," *IEEE Transactions on Communications*, vol. 23, no. 12, pp. 1400–1416, 12 1975. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1092768

[144] M. Kodialam and T. Nandagopal, "Characterizing the capacity region in multi-radio multi-channel wireless mesh networks," in *Proceedings of the 11th annual international conference on Mobile computing and networking - MobiCom '05*. New York, New York, USA: ACM Press, 2005, p. 73. [Online]. Available: http://portal.acm.org/citation.cfm?doid=1080829.1080837

[145] M. Koseoglu and E. Karasan, "Throughput Modeling of Single Hop CSMA Networks with Non-Negligible Propagation Delay," *IEEE Transactions on Communications*, vol. 61, no. 7, pp. 2911–2923, 7 2013. [Online]. Available: http://ieeexplore.ieee.org/document/6516172/

[146] K. Kotiadis and J. Mingers, "Combining PSMs with hard OR methods: the philosophical and practical challenges," *Journal of the Operational Research Society*, vol. 57, no. 7, pp. 856–867, 7 2006. [Online]. Available: https://www.tandfonline.com/doi/full/10.1057/palgrave.jors.2602147

[147] KPMG, "The new business models for operators," 2011.

[148] D. Krajzewicz and G. Hertkorn, "SUMO (Simulation of Urban MObility) An open-source traffic simulation," ... *Symposium on Simulation ...*, pp. 63–68, 2002. [Online]. Available: http://elib.dlr.de/6661/02/dkrajzew_MESM2002.pdf

[149] P. Kulkarni and F. Cao, "Dynamic Sensitivity Control to improve Spatial Reuse in Dense Wireless LANs," in *Proceedings of the 19th ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems - MSWiM '16*. New York, New York, USA: ACM Press, 2016, pp. 323–329. [Online]. Available: http://dl.acm.org/citation.cfm?doid=2988287.2989138

[150] M. Landry, J.-L. Malouin, and M. Oral, "Model validation in operations research," *European Journal of Operational Research*, vol. 14, no. 3, pp. 207–220, 11 1983. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/0377221783902576

[151] R. Laufer and L. Kleinrock, "The Capacity of Wireless CSMA/CA Networks," *IEEE/ACM Transactions on Networking*, vol. 24, no. 3, pp. 1518–1532, 6 2016. [Online]. Available: http://ieeexplore.ieee.org/document/7121036/

[152] L. Laughlin, M. A. Beach, K. A. Morris, and J. L. Haine, "Optimum Single Antenna Full Duplex Using Hybrid Junctions," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 9, pp. 1653–1661, 9 2014. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6832460

[153] A. M. Law, "How to build valid and credible simulation models," in *Proceedings of the 2009 Winter Simulation Conference (WSC)*. IEEE, 12 2009, pp. 24–33. [Online]. Available: http://ieeexplore.ieee.org/document/5429312/

[154] E. Lebhar and Z. Lotker, "Unit disk graph and physical interference model: Putting pieces together," *2009 IEEE International Symposium on Parallel & Distributed Processing*, pp. 1–8, 2009. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5161009

[155] J.-H. Lee, "Self-Interference Cancelation Using Phase Rotation in Full-Duplex Wireless," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 9, pp. 4421–4429, 11 2013. [Online]. Available: http://ieeexplore.ieee.org/document/6517516/

[156] S. Levy, "The Inside Story of the Moto X: The Reason Google Bought Motorola," 2013. [Online]. Available: http://www.wired.com/2013/08/inside-story-of-moto-x/

[157] B. Li, Q. Qu, Z. Yan, and M. Yang, "Survey on OFDMA based MAC protocols for the next generation WLAN," in *2015 IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*. IEEE, 3 2015, pp. 131–135. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=7122542

[158] H. Li, G. Shou, Y. Hu, and Z. Guo, "Mobile Edge Computing: Progress and Challenges," in *2016 4th IEEE International Conference on Mobile Cloud Computing, Services, and Engineering (MobileCloud)*. IEEE, 3 2016, pp. 83–84. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=7474412

[159] W. Liang, Z. Li, H. Zhang, S. Wang, and R. Bie, "Vehicular Ad Hoc Networks: Architectures, Research Issues, Methodologies, Challenges, and Trends," *International Journal of Distributed Sensor Networks*, vol. 11, no. 8, p. 745303, 8 2015. [Online]. Available: http://journals.sagepub.com/doi/10.1155/2015/745303

[160] R. Liao, B. Bellalta, and M. Oliver, "Modelling and Enhancing Full-Duplex MAC for Single-Hop 802.11 Wireless Networks," *IEEE Wireless Communications Letters*, vol. 4,

no. 4, pp. 349–352, 8 2015. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=7069230

[161] S. Liew, C. Kai, J. Leung, and B. Wong, "Back-of-the-envelope computation of throughput distributions in CSMA wireless networks," *IEEE Transactions on Mobile Computing*, vol. 9, no. 9, pp. 1319–1331, 2010. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5467085

[162] W. Liu, X. Wang, W. Zhang, L. Yang, and C. Peng, "Coordinative simulation with SUMO and NS3 for Vehicular Ad Hoc Networks," in *2016 22nd Asia-Pacific Conference on Communications (APCC)*. IEEE, 8 2016, pp. 337–341. [Online]. Available: http://ieeexplore.ieee.org/document/7581471/

[163] U. Mahola and L. Erasmus, "Emerging revenue model structure for mobile industry: The case for traditional and OTT service providers in Sub-Sahara," in *2015 Portland International Conference on Management of Engineering and Technology (PICMET)*. IEEE, 8 2015, pp. 1485–1494. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=7273046

[164] M. W. Maier and E. Rechtin, *The art of systems architecting*. CRC press, 2009.

[165] G. Mao, "Research on wireless multi-hop networks: Current state and challenges," in *2012 International Conference on Computing, Networking and Communications (ICNC)*. IEEE, 1 2012, pp. 593–598. [Online]. Available: http://ieeexplore.ieee.org/document/6167492/

[166] E. Marashi and J. P. Davis, "An argumentation-based method for managing complex issues in design of infrastructural systems," *Reliability Engineering & System Safety*, vol. 91, no. 12, pp. 1535–1545, 2006. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S095183200600038X

[167] ——, "A Systems-Based Approach for Supporting Discourse in Decision Making," *Computer-Aided Civil and Infrastructure Engineering*, vol. 22, no. 7, pp. 511–526, 10 2007. [Online]. Available: http://doi.wiley.com/10.1111/j.1467-8667.2007.00507.x

[168] S. E. Marashi, J. P. Davis, and J. W. Hall, "Combination Methods and Conflict Handling in Evidential Theories," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 16, no. 03, pp. 337–369, 6 2008. [Online]. Available: http://www.worldscientific.com/doi/abs/10.1142/S0218488508005315

[169] A. Maria, "Introduction to Modeling and Simulation," in *Proceedings of the Winter Simulation Conference, 2005.* IEEE, 1997, pp. 16–23. [Online]. Available: http://ieeexplore.ieee.org/document/1574235/

[170] MathWorks, "MATLAB." [Online]. Available: https://uk.mathworks.com/products/matlab.html

[171] N. S. Matloff, "Introduction to discrete-event simulation and the simpy language," *Davis, CA. Dept of Computer Science. University*, pp. 1–33, 2008. [Online]. Available: http://heather.cs.ucdavis.edu/~matloff/156/PLN/DESimIntro.pdf

[172] K. Medepalli and F. A. Tobagi, "Towards Performance Modeling of IEEE 802.11 Based Wireless Networks: A Unified Framework and Its Applications," *Proceedings IEEE INFOCOM 2006. 25TH IEEE International Conference on Computer Communications*, vol. 00, no. c, pp. 1–12, 2006. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4146872

[173] M.-A. Messous, S.-M. Senouci, and H. Sedjelmaci, "Network connectivity and area coverage for UAV fleet mobility model with energy constraint," in *2016 IEEE Wireless Communications and Networking Conference*, vol. 2016-Septe, no. Wcnc. IEEE, 4 2016, pp. 1–6. [Online]. Available: http://ieeexplore.ieee.org/document/7565125/

[174] METIS, "The METIS 2020 Project - Laying the foundation of 5G." [Online]. Available: https://www.metis2020.com/

[175] S. P. Meyn and R. L. Tweedie, *Markov Chains and Stochastic Stability*. Springer-Verlag, 1993. [Online]. Available: http://probability.ca/MT/BOOK.pdf

[176] V. P. Mhatre, K. Papagiannaki, and F. Baccelli, "Interference Mitigation Through Power Control in High Density 802.11 WLANs," in *IEEE INFOCOM 2007 - 26th IEEE International Conference on Computer Communications*. IEEE, 2007, pp. 535–543. [Online]. Available: http://ieeexplore.ieee.org/document/4215651/

[177] J. Mingers, "A classification of the philosophical assumptions of management science methods," *Journal of the Operational Research Society*, vol. 54, no. 6, pp. 559–570, 6 2003. [Online]. Available: http://www.palgrave-journals.com/jors/journal/v54/n6/abs/2601436a.html

[178] ——, "Soft OR comes of age - but not everywhere!" *Omega-International Journal of Management Science*, vol. 39, no. 6, pp. 729–741, 12 2011. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S0305048311000089

[179] J. Mingers and J. Rosenhead, "Problem structuring methods in action," *European Journal of Operational Research*, vol. 152, no. 3, pp. 530–554, 2 2004. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S0377221703000560

[180] Q. T. Minh, K. Nguyen, C. Borcea, and S. Yamada, "On-the-fly establishment of multihop wireless access networks for disaster recovery," *IEEE Communications Magazine*, vol. 52, no. 10, pp. 60–66, 10 2014. [Online]. Available: http://ieeexplore.ieee.org/document/6917403/

[181] R. Mintz, "Systems engineering fundamentals," in *Proceedings of NORTHCON '94*. IEEE, 1994, pp. 316–319. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=643369

[182] D. Miorandi and E. Altman, "Coverage and connectivity of ad hoc networks in presence of channel randomness," in *Proceedings IEEE 24th Annual Joint Conference of the IEEE Computer and Communications Societies.*, vol. 1, no. c. IEEE, 2005, pp. 491–502. [Online]. Available: http://ieeexplore.ieee.org/document/1497917/

[183] D. Miorandi, E. Altman, and G. Alfano, "The Impact of Channel Randomness on Coverage and Connectivity of Ad Hoc and Sensor Networks," *IEEE Transactions on Wireless Communications*, vol. 7, no. 3, pp. 1062–1072, 3 2008. [Online]. Available: http://ieeexplore.ieee.org/document/4471986/

[184] Motorola, "Motorola History Timeline," 2015. [Online]. Available: http://www.motorola.com/us/consumers/about-motorola-us/About_Motorola-History-Timeline/About_Motorola-History-Timeline.html

[185] R. Muñoz and K.-I. Kitayama, "At A Glance: STRAUSS Scalable and efficienT oRchestrAtion of Ethernet services Using Software- defined and flexible optical networkS Project Coordinator: Orchestration of heterogeneous optical networks for Ethernet transport," 2013. [Online]. Available: http://www.ict-strauss.eu/images/STRAUSS_Factsheet_vfinal.pdf

[186] B. Nardelli and E. W. Knightly, "Closed-form throughput expressions for CSMA networks with collisions and hidden terminals," *2012 Proceedings IEEE INFOCOM*, pp. 2309–2317, 2012. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6195618

[187] G. Narlikar, G. Wilfong, and L. Zhang, "Designing Multihop Wireless Backhaul Networks with Delay Guarantees," in *Proceedings IEEE INFOCOM 2006. 25TH IEEE International Conference on Computer Communications*, vol. 00, no. c. IEEE, 2006, pp. 1–12. [Online]. Available: http://ieeexplore.ieee.org/document/4146972/

[188] Nasa, "NASA Systems Engineering Handbook," *Nasa/Sp-2007-6105*, 2007. [Online]. Available: http://ntrs.nasa.gov/

[189] M. Natkaniec and A. Pach, "An analysis of the backoff mechanism used in IEEE 802.11 networks," *Proceedings ISCC 2000. Fifth IEEE Symposium on Computers and Communications*, pp. 444–449, 2000. [Online]. Available: http://ieeexplore.ieee.org/document/860678/

[190] NDFIS, "National Dark Fibre Infrastructure Service." [Online]. Available: http://www.ndfis.org/

[191] P. C. Ng and S. C. Liew, "Throughput Analysis of IEEE802.11 Multi-Hop Ad Hoc Networks," *IEEE/ACM Transactions on Networking*, vol. 15, no. 2, pp. 309–322, 4 2007. [Online]. Available: http://ieeexplore.ieee.org/document/4154761/

[192] NGMN Alliance, "5G White Paper," pp. 1–125, 2015. [Online]. Available: https://www.ngmn.org/uploads/media/NGMN_5G_White_Paper_V1_0_01.pdf

[193] H. Q. Nguyen, F. Baccelli, and D. Kofman, "A Stochastic Geometry Analysis of Dense IEEE 802.11 Networks," *IEEE INFOCOM 2007 - 26th IEEE International Conference on Computer Communications*, pp. 1199–1207, 2007. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4215725

[194] R. S. Nickerson, "Confirmation bias: A ubiquitous phenomenon in many guises." *Review of General Psychology*, vol. 2, no. 2, pp. 175–220, 1998. [Online]. Available: http://doi.apa.org/getdoi.cfm?doi=10.1037/1089-2680.2.2.175

[195] Nokia, "Differentiating your business in a 5G world," 2014. [Online]. Available: http://networks.nokia.com/news-events/insight-newsletter/articles/differentiating-your-business-in-a-5g-world

[196] NS-3, "NS-3," 2016. [Online]. Available: https://www.nsnam.org/

[197] ——, "NS-3: Documentation," 2016. [Online]. Available: https://www.nsnam.org/doxygen/

[198] OPERA, "OPERA - Opportunistic Passive Radar for Non-Cooperative Contextual Sensing." [Online]. Available: http://www.bristol.ac.uk/news/2018/may/opera-project.html

[199] OPNET, "OPNET Technologies - Network Simulator | Riverbed." [Online]. Available: http://www.riverbed.com/gb/products/steelcentral/opnet.html?redirect=opnet

[200] C. A. O'Reilly and M. L. Tushman, "Organizational Ambidexterity: Past, Present, and Future," *Academy of Management Perspectives*, vol. 27, no. 4, pp. 324–338, 11 2013. [Online]. Available: http://amp.aom.org/cgi/doi/10.5465/amp.2013.0025

[201] C. A. O'Reilly, J. B. Harreld, and M. L. Tushman, "Organizational Ambidexterity: IBM and Emerrging Business Opportunities," *California Management Review*, vol. 51, no. 4,

pp. 75–100, 2009. [Online]. Available: http://journals.sagepub.com/doi/abs/10.2307/41166506

[202] A. Osseiran, F. Boccardi, V. Braun, K. Kusume, P. Marsch, M. Maternia, O. Queseth, M. Schellmann, H. Schotten, H. Taoka, H. Tullberg, M. A. Uusitalo, B. Timus, and M. Fallgren, "Scenarios for 5G mobile and wireless communications: the vision of the METIS project," *IEEE Communications Magazine*, vol. 52, no. 5, pp. 26–35, 5 2014. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6815890

[203] M. Ozdemir and A. B. McDonald, "A queuing model of multi-hop wireless ad hoc network with hidden nodes," *The 7th IFIP Int. Conf. Mobile and Wireless Communications Networks (MWCN 2005), Morocco*, p. 19–21, 2005. [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.133.6092&rep=rep1&type=pdf

[204] B. Panzner, W. Zirwas, S. Dierks, M. Lauridsen, P. Mogensen, K. Pajukoski, and D. Miao, "Deployment and implementation strategies for massive MIMO in 5G," in *2014 IEEE Globecom Workshops (GC Wkshps)*. IEEE, 12 2014, pp. 346–351. [Online]. Available: http://ieeexplore.ieee.org/document/7063455/

[205] Q. Qu, B. Li, M. Yang, and Z. Yan, "An OFDMA based concurrent multiuser MAC for upcoming IEEE 802.11ax," in *2015 IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*. IEEE, 3 2015, pp. 136–141. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=7122543

[206] B. Radunovic, D. Gunawardena, P. Key, A. Proutiere, N. Singh, V. Balan, and G. Dejean, "Rethinking Indoor Wireless Mesh Design: Low Power, Low Frequency, Full-Duplex," in *2010 Fifth IEEE Workshop on Wireless Mesh Networks*. IEEE, 6 2010, pp. 1–6. [Online]. Available: http://ieeexplore.ieee.org/document/5507905/

[207] S. Rajagopal, "Power efficiency: The next challenge for multi-gigabit-per-second Wi-Fi," *IEEE Communications Magazine*, vol. 52, no. 11, pp. 40–45, 11 2014. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6957141

[208] R. Ramanathan, "On the performance of ad hoc networks with beamforming antennas," in *Proceedings of the 2nd ACM international symposium on Mobile ad hoc networking & computing - MobiHoc '01*. New York, New York, USA: ACM Press, 2001, p. 95. [Online]. Available: http://portal.acm.org/citation.cfm?doid=501416.501430

[209] T. S. Rappaport, *Wireless communications: principles and practice*. Prentice Hall PTR New Jersey, 1996, vol. 2.

[210] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez, "Millimeter Wave Mobile Communications for 5G Cellular: It Will Work!" *IEEE Access*, vol. 1, pp. 335–349, 2013. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6515173

[211] N. S. Ravindranath, I. Singh, A. Prasad, and V. S. Rao, "Performance Evaluation of IEEE 802.11ac and 802.11n using NS3," *Indian Journal of Science and Technology*, vol. 9, no. 26, 7 2016. [Online]. Available: http://www.indjst.org/index.php/indjst/article/view/93565

[212] A. Reiter, B. Prunster, and T. Zefferer, "Hybrid Mobile Edge Computing: Unleashing the Full Potential of Edge Computing in Mobile Device Use Cases," in *2017 17th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGRID)*. IEEE, 5 2017, pp. 935–944. [Online]. Available: http://ieeexplore.ieee.org/document/7973801/

[213] T. Riihonen and R. Wichman, "Analog and digital self-interference cancellation in full-duplex MIMO-OFDM transceivers with limited resolution in A/D conversion," in *2012 Conference Record of the Forty Sixth Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*. IEEE, 11 2012, pp. 45–49. [Online]. Available: http://ieeexplore.ieee.org/document/6488955/

[214] H. W. J. Rittel and M. M. Webber, "Dilemmas in a general theory of planning," *Policy Sciences*, vol. 4, no. 2, pp. 155–169, 6 1973. [Online]. Available: http://link.springer.com/10.1007/BF01405730

[215] S. Robinson, *Simulation: The Practice of Model Development and Use*, 2003.

[216] J. Rosenhead, "Past, present and future of problem structuring methods," *Journal of the Operational Research Society*, vol. 57, no. 7, pp. 759–765, 7 2006. [Online]. Available: http://link.springer.com/10.1057/palgrave.jors.2602206

[217] ——, "What's the Problem? An Introduction to Problem Structuring Methods," *Interfaces*, vol. 26, no. 6, pp. 117–131, 12 1996. [Online]. Available: http://pubsonline.informs.org/doi/abs/10.1287/inte.26.6.117

[218] E. A. J. A. Rouwette, "Facilitated modelling in strategy development: measuring the impact on communication, consensus and commitment," *Journal of the Operational Research Society*, vol. 62, no. 5, pp. 879–887, 5 2011. [Online]. Available: https://www.tandfonline.com/doi/full/10.1057/jors.2010.78

[219] A. Sabharwal, P. Schniter, D. Guo, D. W. Bliss, S. Rangarajan, and R. Wichman, "In-Band Full-Duplex Wireless: Challenges and Opportunities," *IEEE Journal on Selected*

Areas in Communications, vol. 32, no. 9, pp. 1637–1652, 9 2014. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6832464

[220] A. Sahai, G. Patel, C. Dick, and A. Sabharwal, "On the Impact of Phase Noise on Active Cancelation in Wireless Full-Duplex," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 9, pp. 4494–4510, 11 2013. [Online]. Available: http://ieeexplore.ieee.org/document/6523998/

[221] A. Sahai, G. Patel, and A. Sabharwal, "Pushing the limits of Full-duplex: Design and Real-time Implementation," 7 2011. [Online]. Available: http://arxiv.org/abs/1107.0607

[222] ——, "Asynchronous full-duplex wireless," in *2012 Fourth International Conference on Communication Systems and Networks (COMSNETS 2012)*. IEEE, 1 2012, pp. 1–9. [Online]. Available: http://ieeexplore.ieee.org/document/6151328/

[223] P. Santi and D. Blough, "The critical transmitting range for connectivity in sparse wireless ad hoc networks," *IEEE Transactions on Mobile Computing*, vol. 2, no. 1, pp. 25–39, 1 2003. [Online]. Available: http://ieeexplore.ieee.org/document/1195149/

[224] P. Santi, D. M. Blough, and F. Vainstein, "A probabilistic analysis for the range assignment problem in ad hoc networks," in *Proceedings of the 2nd ACM international symposium on Mobile ad hoc networking & computing - MobiHoc '01*. New York, New York, USA: ACM Press, 2001, p. 212. [Online]. Available: http://portal.acm.org/citation.cfm?doid=501416.501446

[225] R. Sargent, "Verification and validation of simulation models," *Proceedings of the 37th conference on Winter simulation*, p. 130–143, 2005. [Online]. Available: http://dl.acm.org/citation.cfm?id=1162736

[226] N. I. Sarkar and S. A. Halim, "A Review of Simulation of Telecommunication Networks : Simulators , Classification , Comparison , Methodologies , and Recommendations," *Multidisciplinary Journals in Science and Technology, Journal of Selected Areas in Telecommunications (JSAT)*, vol. 2, no. 3, pp. 10–17, 2011. [Online]. Available: http://www.cyberjournals.com/Papers/Mar2011/02.pdf

[227] A. Sayenko, M. Zolotukhin, and T. Hamalainen, "Multi-hop relays for high frequency next generation wireless systems," in *2017 24th International Conference on Telecommunications (ICT)*. IEEE, 5 2017, pp. 1–7. [Online]. Available: http://ieeexplore.ieee.org/document/7998236/

[228] W. Schacherbauer, T. Ostertag, C. Ruppel, A. Springer, and R. Weigel, "An Interference Cancellation Technique for the Use in Multiband Software Radio Frontend Design,"

in *30th European Microwave Conference, 2000*. IEEE, 10 2000, pp. 1–4. [Online]. Available: http://ieeexplore.ieee.org/document/4139855/

[229] T. Schriber and D. Brunner, "Inside simulation software: how it works and why it matters," in *Proceedings of Winter Simulation Conference*, no. 1998. IEEE, 2007, pp. 45–54. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=717072

[230] P. M. Senge, "The Fifth Discipline," *Measuring Business Excellence*, vol. 1, no. 3, pp. 46–51, 3 1997. [Online]. Available: http://www.emeraldinsight.com/doi/10.1108/eb025496

[231] G. Shen, J. Liu, D. Wang, J. Wang, and S. Jin, "Multi-hop relay for next-generation wireless access networks," *Bell Labs Technical Journal*, vol. 13, no. 4, pp. 175–193, 2 2009. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6769034

[232] F. Simon, "Uncertainty and Imprecision: Modelling and Analysis," *The Journal of the Operational Research Society*, vol. 46, no. 1, pp. 70–79, 1995. [Online]. Available: http://www.jstor.org/stable/2583837

[233] M. Soroushnejad and E. Geraniotis, "Probability of capture and rejection of primary multiple-access interference in spread-spectrum networks," *IEEE Transactions on Communications*, vol. 39, no. 6, pp. 986–994, 6 1991. [Online]. Available: http://ieeexplore.ieee.org/document/87188/

[234] SPHERE, "SPHERE : An EPSRC Interdisciplinary Research Collaboration (IRC)," 2015. [Online]. Available: http://www.irc-sphere.ac.uk/about

[235] J. Sturgeon, "Where Nortel went wrong | Financial Post," 2012. [Online]. Available: http://business.financialpost.com/fp-tech-desk/where-nortel-went-wrong

[236] V. G. Subramanian and D. J. Leith, "Convexity Conditions for 802.11 WLANs," pp. 1–5, 6 2012. [Online]. Available: http://arxiv.org/abs/1206.3120

[237] ——, "On the Rate Region of CSMA/CA WLANs," *IEEE Transactions on Information Theory*, vol. 59, no. 6, pp. 3932–3938, 6 2013. [Online]. Available: http://ieeexplore.ieee.org/document/6482633/

[238] J. Sztrik, "Basic queueing theory," *University of Debrecen, Faculty of Informatics*, vol. 193, 2012. [Online]. Available: https://pdfs.semanticscholar.org/848f/a1f48ad9d3edb24b05667f15cfc633eb8f69.pdf

[239] M. Takai, J. Martin, and R. Bagrodia, "Effects of wireless physical layer modeling in mobile ad hoc networks," in *Proceedings of the 2nd ACM international*

symposium on Mobile ad hoc networking & computing - MobiHoc '01, ser. MobiHoc '01. New York, New York, USA: ACM Press, 2001, p. 87. [Online]. Available: http://doi.acm.org/10.1145/501426.501429

[240] TALON, "TALON." [Online]. Available: http://www.talon.world/

[241] K. Tamaki, A. Raptino H., Y. Sugiyama, M. Bandai, S. Saruwatari, and T. Watanabe, "Full Duplex Media Access Control for Wireless Multi-Hop Networks," in *2013 IEEE 77th Vehicular Technology Conference (VTC Spring)*. IEEE, 6 2013, pp. 1–5. [Online]. Available: http://ieeexplore.ieee.org/document/6692573/

[242] A. Tang, C. Florens, and S. Low, "An empirical study on the connectivity of ad hoc networks," in *2003 IEEE Aerospace Conference Proceedings (Cat. No.03TH8652)*, vol. 3, no. 4. IEEE, 2004, pp. 1333–3. [Online]. Available: http://ieeexplore.ieee.org/document/1235249/

[243] L. Tassiulas and A. Ephremides, "Jointly optimal routing and scheduling in packet ratio networks," *IEEE Transactions on Information Theory*, vol. 38, no. 1, pp. 165–168, 1992. [Online]. Available: http://ieeexplore.ieee.org/document/108264/

[244] ——, "Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks," *IEEE Transactions on Automatic Control*, vol. 37, no. 12, pp. 1936–1948, 1992. [Online]. Available: http://ieeexplore.ieee.org/document/182479/

[245] P. Taylor, "AT&T leads the charge into 'connected' cars - FT.com," 2014. [Online]. Available: http://www.ft.com/cms/s/0/5d1ea1c6-7b75-11e3-a2da-00144feabdc0.html#axzz3qRXQI900

[246] The-Guardian, "Facebook Buys Oculus," 2014. [Online]. Available: http://www.theguardian.com/technology/2014/jul/22/facebook-oculus-rift-acquisition-virtual-reality(accessed14/01/2015

[247] The-Wearable-Clinic, "The Wearable Clinic: Connecting Health, Self and Care | HeRC." [Online]. Available: https://www.herc.ac.uk/research_project/wearable-clinic-connecting-health-self-care/

[248] C. Thorpe and L. Murphy, "A Survey of Adaptive Carrier Sensing Mechanisms for IEEE 802.11 Wireless Networks," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 3, pp. 1266–1293, 1 2014. [Online]. Available: http://ieeexplore.ieee.org/document/6780905/

[249] A. Ting, A. Ghaleb, K. H. Kwong, K.-C. Lim, D. Chieng, and H.-S. Lim, "Throughput Analysis of IEEE802.11n using OPNET," in *IET International*

Conference on Wireless Communications and Applications (ICWCA 2012). Institution of Engineering and Technology, 2012, pp. 014–014. [Online]. Available: http://digital-library.theiet.org/content/conferences/10.1049/cp.2012.2068

[250] I. Tinnirello, G. Bianchi, and Yang Xiao, "Refinements on IEEE 802.11 Distributed Coordination Function Modeling Approaches," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 3, pp. 1055–1067, 3 2010. [Online]. Available: http://ieeexplore.ieee.org/document/5191039/

[251] TOUCAN, "TOUCAN," 2015. [Online]. Available: http://www.toucan-network.ac.uk/

[252] S. Toumpis and A. Goldsmith, "Capacity regions for wireless ad hoc networks," *IEEE Transactions on Wireless Communications*, vol. 24, no. 5, pp. 736–748, 5 2003. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1210740

[253] J. T. Townsend, J. R. Busemeyer, J. Vandekerckhove, D. Matzke, and E.-J. Wagenmakers, "Model Comparison and the Principle of Parsimony," *The Oxford Handbook of Computational and Mathematical Psychology*, 2015. [Online]. Available: http://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780199957996.001.0001/oxfordhb-9780199957996-e-14

[254] UBIMON, "UBIMON - Ubiquitous Monitoring Environment for Wearable and Implantable Sensors." [Online]. Available: https://www.doc.ic.ac.uk/vip/ubimon/home/

[255] UK-Autodrive, "The UK Autodrive project." [Online]. Available: http://www.ukautodrive.com/

[256] UMTSWORLD, "The history of UMTS and 3G development," 2006. [Online]. Available: http://www.umtsworld.com/umts/history.htm

[257] University-of-Bristol-&-Innovate-UK, "Venturer," 2014. [Online]. Available: http://www.bristol.ac.uk/news/2015/february/venturer.html

[258] J. Utterback, *Mastering the Dynamics of Innovation*. Boston: Harvard Business School Press, 1994.

[259] P. M. Van De Ven, S. C. Borst, J. S. H. Van Leeuwaarden, and A. Proutire, "Insensitivity and stability of random-access networks," *Performance Evaluation*, vol. 67, no. 11, pp. 1230–1242, 2010. [Online]. Available: http://dx.doi.org/10.1016/j.peva.2010.08.011

[260] B. Van Veen and K. Buckley, "Beamforming: a versatile approach to spatial filtering," *IEEE ASSP Magazine*, vol. 5, no. 2, pp. 4–24, 4 1988. [Online]. Available: http://ieeexplore.ieee.org/document/665/

[261] Volvo, "IntelliSafe Autopilot | Volvo Cars," 2015. [Online]. Available: http://www.volvocars.com/intl/about/our-innovation-brands/intellisafe/intellisafe-autopilot

[262] F. W. Vook, A. Ghosh, and T. A. Thomas, "MIMO and beamforming solutions for 5G technology," *2014 IEEE MTT-S International Microwave Symposium (IMS2014)*, pp. 1–4, 6 2014. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6848613

[263] D. H. T. Walker and L. Bourne, "Visualising and Mapping Stakeholder Influence," *Management Decision*, vol. 43, no. 5, 2005. [Online]. Available: http://www.emeraldinsight.com/doi/abs/10.1108/00251740510597680

[264] D. H. T. Walker, L. M. Bourne, and A. Shelley, "Influence, stakeholder mapping and visualization," *Construction Management and Economics*, vol. 26, no. 6, pp. 645–658, 6 2008. [Online]. Available: http://www.tandfonline.com/doi/abs/10.1080/01446190701882390

[265] S. Wang, V. Venkateswaran, and X. Zhang, "Exploring full-duplex gains in multi-cell wireless networks: A spatial stochastic framework," in *2015 IEEE Conference on Computer Communications (INFOCOM)*. IEEE, 4 2015, pp. 855–863. [Online]. Available: http://ieeexplore.ieee.org/document/7218456/

[266] X. Wang and K. Kar, "Throughput modelling and fairness issues in CSMA/CA based ad-hoc networks," in *Proceedings IEEE 24th Annual Joint Conference of the IEEE Computer and Communications Societies.*, vol. 1, no. c. IEEE, 2005, pp. 23–34. [Online]. Available: http://ieeexplore.ieee.org/document/1497875/

[267] X. Wang, H. Huang, and T. Hwang, "On the Capacity Gain from Full Duplex Communications in A Large Scale Wireless Network," *IEEE Transactions on Mobile Computing*, vol. 1233, no. c, pp. 1–1, 2015. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=7300430

[268] D. J. Watts and S. H. Strogatz, "Collective dynamics of small world networks," *Nature*, vol. 393, no. 6684, pp. 440–442, 6 1998. [Online]. Available: http://dx.doi.org/10.1038/30918

[269] L. White, "Understanding problem structuring methods interventions," *European Journal of Operational Research*, vol. 199, no. 3, pp. 823–833, 12 2009. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S0377221709002021

[270] B. Whitworth, "Socio-technical systems," *Encyclopedia of human computer interaction*, pp. 533–541, 2006. [Online]. Available: http://brianwhitworth.com/hci-sts.pdf

[271] Wirelesshistoryfoundation, "Wireless History Timeline | Wireless History Foundation," 2015. [Online]. Available: http://www.wirelesshistoryfoundation.org/wireless-history-project/wireless-history-timeline

[272] H. Wu, Y. Peng, K. Long, and S. Cheng, "A simple model of IEEE 802.11 wireless LAN," in *2001 International Conferences on Info-Tech and Info-Net. Proceedings (Cat. No.01EX479)*, vol. 2. IEEE, 2001, pp. 514–519. [Online]. Available: http://ieeexplore.ieee.org/document/983630/

[273] X. Xie and X. Zhang, "Does full-duplex double the capacity of wireless networks?" in *IEEE INFOCOM 2014 - IEEE Conference on Computer Communications*, no. 978. IEEE, 4 2014, pp. 253–261. [Online]. Available: http://ieeexplore.ieee.org/document/6847946/

[274] K. Xu, M. Gerla, and S. Bae, "How effective is the IEEE 802.11 RTS/CTS handshake in ad hoc networks," in *Global Telecommunications Conference, 2002. GLOBECOM '02. IEEE*, vol. 1. IEEE, 2002, pp. 72–76. [Online]. Available: http://ieeexplore.ieee.org/document/1188044/

[275] T. Yang, G. Mao, and W. Zhang, "Connectivity of Large-Scale CSMA Networks," *IEEE Transactions on Wireless Communications*, vol. 11, no. 6, pp. 2266–2275, 6 2012. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6189002

[276] Y. Yang, B. Chen, K. Srinivasan, and N. B. Shroff, "Characterizing the achievable throughput in wireless networks with two active RF chains," in *IEEE INFOCOM 2014 - IEEE Conference on Computer Communications*, vol. 1, no. c. IEEE, 4 2014, pp. 262–270. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6847947

[277] V. Yazıcı, U. Kozat, and M. Sunay, "A new control plane for 5G network architecture with a case study on unified handoff, mobility, and routing management," *IEEE Communications Magazine*, vol. 52, no. 11, pp. 76–85, 11 2014. [Online]. Available: http://ieeexplore.ieee.org/document/6957146/

[278] M. Yearworth, J. W. Singer, R. Adcock, D. Hybertson, M. Singer, G. Chroust, and K. J. Kijima, "Systems Engineering in a Context of Systemic Cooperation (SCOOPs): Development and Implications," *Procedia Computer Science*, vol. 44, pp. 214–223, 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1877050915002847

[279] M. Yearworth and L. White, "The non-codified use of problem structuring methods and the need for a generic constitutive definition," *European Journal of Operational Research*, vol. 237, no. 3, pp. 932–945, 9 2014. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S0377221714001301

[280] L.-H. Yen and Y.-M. Cheng, "Clustering coefficient of wireless ad hoc networks and the quantity of hidden terminals," *IEEE Communications Letters*, vol. 9, no. 3, pp. 234–236, 3 2005. [Online]. Available: http://ieeexplore.ieee.org/document/1411017/

[281] X. Zhang, W. Cheng, and H. Zhang, "Full-duplex transmission in PHY and MAC layers for 5G mobile wireless networks," *IEEE Wireless Communications*, vol. 22, no. 5, pp. 112–121, 10 2015. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=7306545

[282] Z. Zhong, P. Kulkarni, F. Cao, Z. Fan, and S. Armour, "Issues and challenges in dense WiFi networks," in *2015 International Wireless Communications and Mobile Computing Conference (IWCMC)*. IEEE, 8 2015, pp. 947–951. [Online]. Available: http://ieeexplore.ieee.org/document/7289210/

[283] C. Zhu, O. Yang, J. Aweya, M. Ouellette, and D. Montuno, "A comparison of active queue management algorithms using the OPNET Modeler," *IEEE Communications Magazine*, vol. 40, no. 6, pp. 158–167, 6 2002. [Online]. Available: http://ieeexplore.ieee.org/document/1007422/

[284] M. Zorzi and S. Pupolin, "Optimum transmission ranges in multihop packet radio networks in the presence of fading," *IEEE Transactions on Communications*, vol. 43, no. 7, pp. 2201–2205, 7 1995. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=392962